

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR



TRABAJO FIN DE MÁSTER

Estudio de Asistentes Virtuales en el Entorno de la Diversidad Funcional Intelectual

**Máster Universitario en Investigación e Innovación en
TIC (I2-TIC)**

Autor: TORRES VALVERDE, Leonardo

Tutor: MONTORO MANRIQUE, Germán

Departamento de Ingeniería Informática

FECHA: Julio, 2019

Resumen

La diversidad funcional intelectual se refiere a limitaciones cognitivas que influyen en el funcionamiento mental y comportamiento adaptativo de una persona. Quienes están considerados dentro de este entorno presentan dificultades en su capacidad de socializar, memorizar, tolerancia reducida a equivocaciones y problemas de comunicación debido a un limitado desarrollo del lenguaje. Este último, genera una barrera con lo cual puede causar un rápido abandono en tareas en donde es necesario la constancia y repetición.

Con la aparición de asistentes de voz los cuales son agentes de software dotados de inteligencia artificial, en donde la principal forma de interacción con los mismos es mediante el uso de la voz, han llevado a la aparición de nuevos dispositivos como es el caso de pantallas y altavoces inteligentes. Marcas como Amazon, Apple, Google, Samsung y Microsoft han lanzado altavoces los cuales contienen a sus asistentes de voz propios. Dichos dispositivos permiten realizar tareas que usualmente se realizan haciendo uso de interfaces gráficas, para lo cual hacen uso de una palabra con lo cual comienzan a percibir peticiones.

El presente estudio se centra en el uso de asistentes de voz comerciales sobre altavoces inteligentes siendo usados por personas con diversidad funcional intelectual. Con lo cual se realizaron sesiones con un total de 26 participantes divididos en 3 grupos homogéneos, los cuales realizaron diversos tipos de tareas con los asistentes de voz sobre altavoces inteligentes. Al concluir el estudio y analizar los datos obtenidos se evidencio las limitaciones que presentan al ser usados por personas con diversidad funcional intelectual puesto que no han sido valoradas las condiciones de su entorno. En adición, se proponen unas guías de diseño las cuales buscan eliminar las barreras con las que se han encontrado los participantes dentro del estudio.

Palabras Clave

Diversidad Funcional Intelectual, Discapacidad Intelectual, Altavoz Inteligente, Asistentes de Voz

Abstract

Intellectual disability refers to cognitive limitations that influence a person's mental functioning and adaptive behavior. Those who are considered within this environment present difficulties in their ability to socialize, memorize, reduced tolerance for mistakes, and communication problems due to limited language development. The latter generates a barrier which can cause a rapid abandonment in tasks where constancy and repetition is necessary.

With the appearance of voice assistants which are software agents equipped with artificial intelligence, where the main form of interaction with them is through the use of voice, have led to the emergence of new devices such as smart screens and smart loudspeakers. Brands such as Amazon, Apple, Google, Samsung and Microsoft have launched speakers which contain their own voice assistants. These devices allow tasks that are usually performed using graphical interfaces, for which they make use of a word with which they begin to perceive requests.

This study focuses on the use of commercial voice assistants on smart speakers being used by people with intellectual disability. Thus, sessions were conducted with a total of 26 participants divided into 3 homogeneous groups, which performed various types of tasks with voice assistants. Upon concluding the study and analyzing the data obtained, the limitations of their use by people with intellectual disability were evident, since the conditions of their environment have not been valued. In addition, some design guides are proposed which seek to eliminate the barriers that have been encountered by participants within the study.

Keywords

Intellectual Disability, Cognitive Disability, Smart Speaker, Voice Assistant.

Agradecimientos

Agradezco principalmente a mis padres Fernando y Rosa por la oportunidad que me han brindado para cumplir este logro, por el apoyo incondicional y consejos brindados a lo largo de mi vida que, en muchas ocasiones me han costado entender. De igual manera agradezco el apoyo de mis hermanos Fer, Heinz y Alvaro que siempre están pendientes de mí, que por difícil que sea de aceptar, la distancia es en donde uno valora su cariño y principalmente su compañía. A mi familia por sus palabras de aliento y sus buenos deseos.

Agradezco también a mis tutores Germán y Javier quienes con su gran experiencia han sabido guiarme en las distintas etapas realizadas en el trabajo. Muchísimas gracias por el apoyo y confianza brindada.

Agradezco a mi amigo Santiago, quien ha sabido apoyarme en momentos difíciles durante este tiempo en Madrid.

Finalmente agradezco a quienes de alguna manera me han sabido apoyar en todo este tiempo lejos de casa y a quienes he conocido en esta maravillosa ciudad.

Índice General

Resumen.....	III
Abstract.....	V
Agradecimientos	VII
Índice General.....	IX
Índice de Figuras.....	XII
Índice de Tablas	XIII
1 Introducción.....	1
1.1 Diversidad Funcional Intelectual	1
1.2 Asistentes de voz comerciales.....	2
1.3 Motivación	2
1.4 Objetivos	2
2 Estado del Arte	3
3 Análisis técnico	7
3.1 Asistentes de voz.....	7
3.1.1 Siri.....	7
3.1.2 Google Assistant	7
3.1.3 Alexa.....	7
3.2 Altavoces inteligentes	8
3.2.1 Homepod.....	8
3.2.2 Google Home	9
3.2.3 Amazon Echo.....	10
3.3 Funcionalidades comunes	10
3.4 Dispositivos adicionales.....	12
3.4.1 Enchufe Koogeek P1	12
3.4.2 Ventilador	12
3.4.3 Phillips Hue Bridge.....	13
3.4.4 Phillips Hue Go.....	13
4 Metodología.....	15
4.1 Procedimiento.....	15
4.2 Tareas	15
4.2.1 Grupo de tareas 1	16
4.2.2 Grupo de tareas 2	16
4.2.3 Grupo de tareas 3	17
4.3 Sesiones.....	17
4.3.1 Sesión de entrenamiento	17

4.3.2	Sesión de estudio	17
4.3.3	Duración de las sesiones	18
4.3.4	Protocolo de sesiones	18
4.4	Análisis de Resultados	18
4.4.1	Frases de activación	19
4.4.2	Tareas realizadas	19
4.4.2.1	Formulación de la tarea	19
4.4.2.2	Respuesta de la tarea	19
5	Evaluación	21
5.1	Participantes	21
5.1.1	Participantes ALENTA	21
5.1.2	Participantes con perfil tecnológico	21
5.2	Materiales	21
5.2.1	Altavoces inteligentes	21
5.2.2	Configuraciones altavoces inteligentes	23
5.2.3	Videograbadora	23
5.2.4	Configuración del entorno de evaluación	23
5.2.5	Pictogramas	23
5.3	Sesiones ALENTA	24
5.3.1	Sesión Entrenamiento	24
5.3.2	Sesiones de Prueba	25
5.3.2.1	Inicio de sesión	25
5.3.2.2	Durante la sesión	25
5.3.3	Entrevista realizada a educadores de ALENTA	27
5.4	Sesiones de participantes con perfil tecnológico	28
5.4.1	Inicio de cada sesión	28
5.4.2	Durante la sesión	28
5.4.3	Entrevista al finalizar la sesión	28
6	Resultados	29
6.1	Frase de activación	29
6.1.1	Activación usando “Oye Siri”	30
6.1.2	Activación usando “Ok Google”	30
6.1.3	Activación usando “Alexa”	31
6.1.4	Comparación entre comandos de activación	32
6.2	Tareas realizadas	33
6.2.1	Tareas realizadas con Siri	33
6.2.2	Tareas realizadas con Google	34

6.2.3	Tareas realizadas con Alexa	36
7	Discusión	39
7.1	Frase de activación	39
7.2	Tareas realizadas	39
7.3	Percepción de activación	40
7.4	Guías de diseño propuesta	40
8	Conclusiones y Trabajo Futuro	41
	Referencias.....	43
	Anexos	45
	Anexo A. Pictogramas usados en las sesiones.....	45
	Anexo B. Preguntas de entrevista.....	48

Índice de Figuras

Figura 3.1. Altavoz Inteligente HomePod	8
Figura 3.2. Altavoz Inteligente Google Home.....	9
Figura 3.3. Altavoz Inteligente Amazon Echo	10
Figura 3.4. Enchufe Koogeek P1	12
Figura 3.5. Ventilador a usarse con enchufe.....	12
Figura 3.6. Dispositivo Phillips Hue Bridge.....	13
Figura 3.7. Lámpara Phillips Go.....	13
Figura 4.1. Procedimiento usando Latin Square.....	16
Figura 5.1. Pictogramas de altavoces inteligentes	23
Figura 6.1: Intentos realizados con Siri	30
Figura 6.2: Intentos realizados con Google	31
Figura 6.3: Intentos realizados con Alexa	31
Figura 6.4: Comparativa entre frases de activación.....	32
Figura 6.5: Primer grupo de tareas realizadas con Siri	33
Figura 6.6 Segundo grupo de tareas realizadas con Siri	34
Figura 6.7: Primer grupo de tareas realizadas con Google	35
Figura 6.8: Segundo grupo de tareas realizadas con Google	35
Figura 6.9: Primer grupo de tareas realizadas con Alexa	36
Figura 6.10: Segundo grupo de tareas realizadas con Alexa	36
Figura 8.1: Pictograma de encender el ventilador	45
Figura 8.2: Pictograma Apaga el ventilador	45
Figura 8.3: Pictograma Enciende la luz	45
Figura 8.4: Pictograma Apaga la luz	46
Figura 8.5: Pictograma de pregunta ¿Qué tiempo hace?	46
Figura 8.6: Pictograma Cuéntame un cuento.....	46
Figura 8.7: Pictograma Cántame una canción	47

Índice de Tablas

Tabla 1. Palabras clave usadas.....	3
Tabla 2. Características de HomePod	8
Tabla 3. Características de Google Home	9
Tabla 4. Características de Amazon Echo.	10
Tabla 5. Altavoces usados en la evaluación	21
Tabla 6: Información de participantes de ALENTA	22
Tabla 7: Cronograma de sesiones en ALENTA	25
Tabla 8: Número de tareas realizadas por altavoz.	29

1 Introducción

En este trabajo se presenta un estudio de asistentes virtuales comerciales sobre altavoces inteligente usados por personas con diversidad funcional intelectual. La finalidad de este estudio es la de conocer las ventajas y los inconvenientes que presentan estos dispositivos puesto que son una alternativa a las interfaces de usuario gráficas. En el estudio, los participantes realizaron una serie de tareas, haciendo uso de los altavoces inteligentes de Apple, Amazon y Google con lo cual se evidencio las barreras que existen entre estos dispositivos y la interacción, además del entendimiento y percepción de un asistente de voz. Para iniciar con el presente estudio se comienza con una definición de diversidad funcional intelectual y asistentes de voz que sirven como entrada para abordar el trabajo realizado.

1.1 Diversidad Funcional Intelectual

La diversidad funcional intelectual es un término el cual se refiere a limitaciones significativas en el funcionamiento mental y comportamiento adaptativo de una persona, las cuales tienen su origen antes de los 18 años. Las causas pueden deberse a factores biológicos como condiciones genéticas, problemas durante el embarazo, al nacer, problemas de salud o factores del entorno como vulnerabilidad social o del entorno [1]. Su severidad va clasificada en cuatro categorías (leve, moderada, severa y profunda) [2]. Dependiendo de la severidad que presenten, requieren la ayuda para realizar tareas básicas y en casos extremos, la presencia continua de una persona que lo asista es necesaria. Refiriéndose a la diversidad funcional intelectual puede incluir personas con trastorno del espectro autista en sus diferentes niveles, síndrome de Down, trastorno del déficit de atención e hiperactividad entre otras.

Sus capacidades de memorización y aprendizaje son significativamente bajas. Tienen una deficiencia en relacionar información y no usan estrategias de retención apropiadas para el aprendizaje y memorización. Presentan problemas de autorregulación los cuales afectan las habilidades organizacionales y el control de los procesos de aprendizaje. Tienden a distraerse con facilidad o prestar atención a cuestiones irrelevantes. En el ámbito interpersonal presentan dificultades distinguiendo situaciones sociales puesto que no entienden o no filtran la información importante lo que conlleva a un comportamiento impulsivo o socialmente incorrecto.

Pueden presentar retrasos en el habla, comprensión del lenguaje y dificultades de formulación, estos problemas están generalmente asociados en el retraso del desarrollo del lenguaje. Pueden presentar funcionamiento retrasado en aspectos pragmáticos del lenguaje como tomar turnos para hablar o seleccionar temas apropiados de conversación.

Por otro lado, las experiencias pasadas en las que hayan fallado o que hayan generado ansiedad pueden hacer que parezcan tener menos objetivos o falta de motivación. Además, es menos probable que sean proactivos al realizar alguna actividad. Pueden presentar discapacidades motoras, visuales, auditivas y problemas de salud. Además, existe una relación entre la severidad de diversidad funcional intelectual y las diferencias físicas de los individuos [1].

1.2 Asistentes de voz comerciales

Un asistente virtual es un agente de software que mediante el uso de reconocimiento de voz e inteligencia artificial crean una interfaz conversacional, la cual permite la realización de tareas que, en un principio fueron diseñadas para ser realizadas a través de interfaces en donde el ingreso de información se realiza por medio de botones o entradas táctiles. Hoy en día, tareas como manipulación de aplicaciones de software, interacción con dispositivos o búsquedas de internet son capaces de ser realizadas por asistentes virtuales. La activación de los mismos se realiza mediante una palabra clave con la cual comienzan a escuchar y desde ese punto reciben órdenes. La aparición de los mismos remonta al año 2011, con la aparición del asistente de voz de Apple llamado Siri, el cual en un principio estaba disponible en teléfonos móviles [3]. Hoy por hoy, estos asistentes se encuentran disponibles en ordenadores, dispositivos móviles, altavoces y relojes inteligentes entre otros.

Algunos de los asistentes virtuales más conocidos son Siri¹ de Apple, Google Assistant² de Google, Alexa de Amazon, Bixby³ de Samsung, Cortana⁴ de Microsoft. Los cuales, dependiendo del dispositivo en el que se encuentren, ofrecen información de manera distinta ya sea por voz, grafica o una combinación de ambas. Recientemente, el uso de los asistentes de voz sobre altavoces inteligentes ha ido en aumento y con ello la aparición de servicios que permiten interactuar con otros dispositivos y realizar tareas comunes que faciliten a los usuarios finales [3].

1.3 Motivación

La motivación para la realización del presente trabajo fue conocer las limitaciones y ventajas que tienen los asistentes de voz sobre altavoces inteligentes al ser usados por personas con diversidad funcional intelectual ya que, la interacción común que tienen con dispositivos tecnológicos es a través de una interfaz en donde ingresan información de manera táctil y reciben respuestas graficas o vocales. Además, es interesante conocer la naturalidad de interactuar con los altavoces y su reacción a las respuestas obtenidas [4][5].

1.4 Objetivos

Objetivo General

- Evaluar las ventajas y limitaciones de los asistentes de voz comerciales al usarse por personas con diversidad funcional intelectual

Objetivos Específicos

- Proponer una metodología de evaluación
- Formular tareas a realizarse por los asistentes de voz
- Realizar un piloto con participantes en el entorno de la diversidad funcional intelectual
- Analizar la información recolectada
- Evidenciar hallazgos de la evaluación realizada

¹ Asistente de voz Siri, fuente: <https://www.apple.com/es/siri/>

² Asistente de voz Google Assistant, fuente: https://assistant.google.com/intl/es_es/

³ Asistente de voz Bixby, fuente: <https://www.samsung.com/es/apps/bixby/>

⁴ Asistente de voz Cortana, fuente: <https://www.microsoft.com/es-es/windows/cortana>

2 Estado del Arte

Se realizó una revisión de la bibliografía sobre asistentes de voz usados por personas con diversidad funcional intelectual. La búsqueda de información se realizó sobre las bases de datos de Scopus, IEEE Xplore y ACM Digital Library puesto que cuentan con artículos referentes al tema y son versátiles en los criterios de búsqueda. Por otro lado, las palabras claves usadas dentro de las búsquedas son las mostradas en la Tabla 1.

Concepto	Sinónimos y términos alternativos
Asistente de voz	Voice assistant
	Virtual assistant
	Voice interface
Discapacidad Intelectual	Cognitive disabilities
	Intellectual disabilities
	Cognitive impairment

Tabla 1. Palabras clave usadas

Criterios de inclusión y exclusión

Se aplicaron criterios de exclusión en la búsqueda en donde se indicó que la fecha de publicación fuera menor al año 2011 y que estuvieran escritos en otros idiomas que no sea el inglés. Mientras que para los criterios de inclusión se consideraron aquellos que hayan propuesto aplicaciones y/o en los cuales haya existido interacción entre sistemas propuestos o comerciales enfocados a personas con diversidad intelectual funcional.

Al haber realizado los pasos antes mencionados se procedió a la revisión de artículos que cumplieren con los criterios antes mencionados y con lo cual se obtuvo la siguiente información:

En el estudio dirigido por Baldauf y su equipo [5], realizaron una entrevista grupal a 5 personas con deterioro cognitivo sobre si harían uso de interfaces conversacionales. Los resultados de dicha entrevista mostraron que, los participantes consideran a sus dispositivos móviles como su compañía, puesto que les ayudan en su vida diaria, sin embargo, por problemas de motricidad y aumento del deterioro cognitivo se les dificulta interactuar con los mismos. Otros problemas reportados están enfocados a sitios web complejos, máquinas de ventas de tickets, entre otros. La entrevista mostro que 2 de los participantes tenían conocimiento sobre los asistentes de voz y los usaban para enviar mensajes, informarse del clima o como un compañero. El resto de los participantes no habían hecho uso de este tipo de interfaz, sin embargo, identificaron un posible obstáculo cuando los asistentes no entendían las palabras o frases dichas y que el lenguaje a ser usado debía ser conocido por las personas que padecen alguna discapacidad cognitiva. Finalmente, los participantes aconsejan que el asistente de voz actué como mediador para interactuar con otros dispositivos.

Con relación al trabajo de investigación se encuentran trabajos que han utilizado asistentes de voz y que han tenido la interacción de personas con diversidad funcional intelectual como es el caso de [6], para este estudio, se planteó el desarrollo de una aplicación, la cual permitía ingresar rutinas a realizarse a diario por personas con discapacidades cognitivas para luego ser utilizadas por el asistente, el cual le haría saber lo que debe hacer de forma hablada. La aplicación desarrollada tiene integración con el asistente de voz de Amazon, Alexa, de modo que, quien utilice la aplicación puede hacer uso de Alexa para saber cómo realizar alguna acción la cual no recuerde como hacerla.

Wallace T. y Morris J. [7] realizaron un estudio de usabilidad para identificar barreras de interacción entre altavoces inteligentes y personas con discapacidades cognitivas. Los altavoces de Amazon y Google fueron usados en el experimento, el cual, partía desde la configuración de los dispositivos a la interacción con los mismos. Se contó con la participación de 8 militares retirados que han sufrido daño cerebral o desorden de estrés postraumático. Inicialmente existieron problemas al configurar los dispositivos debido a la falta de instrucciones escritas por lo que recomiendan una descripción por escrito que muestre los posibles mensajes de error. Luego de interactuar con los dispositivos, los participantes reportaron que el altavoz Amazon Echo es más intuitivo y fácil de instalar a comparación del dispositivo Google Home, sin embargo, el asistente sobre este dispositivo tenía mayor entendimiento de los comandos de voz usados. Finalmente, los participantes reportaron una mayor preferencia por el asistente de voz de Amazon puesto que la interacción es más natural al tener que llamar al dispositivo como a una persona más no como a un robot.

Balasuriya y su equipo [8], realizaron una evaluación de interacción de personas con diversidad funcional intelectual haciendo uso de asistentes de voz, para lo cual se contó con la participación de 18 personas. Los asistentes usados fueron SIRI y Google Assistant. Al realizar las pruebas en dicho estudio, se evidenciaron varios problemas de interacción. Por un lado, problemas de interacción con la interfaz gráfica de los dispositivos en donde se encontraban alojados los asistentes, puesto que podían fallar al escribir en la interfaz.

Así mismo al pronunciar frases largas o hacer pausas, el asistente no comprendía la orden, además de la dificultad de entendimiento por la forma de pronunciación de los participantes. Aunque existían problemas de entendimiento, el estudio mostro que los participantes preferían interactuar de forma hablada puesto que era más común el uso de lenguaje natural, además de que, si no conocían como se escribía alguna palabra, esto se evitaba al hablar.

Myers y su equipo[9], realizaron un estudio de como los usuarios interactúan con interfaces de usuario activadas por voz. Hicieron uso de “DiscoverCal”, una aplicación para administrar el calendario, que hace uso de la herramienta de procesamiento de lenguaje natural DialogFlow de Google. La forma principal de interacción con la aplicación es a través de la voz. Dentro del experimento los participantes tenían que crear, modificar, eliminar e invitar a otros participantes a eventos dentro del calendario. Los resultados revelaron algunos problemas de interacción. El más común se producía debido a la interpretación errónea del comando causando que ejecute una acción equivocada. Otro de los errores encontrados fue la estructura de la petición puesto que el asistente no soporta la petición realizada. Adicionalmente el estudio revelo que los usuarios variaban la forma de interacción con el sistema para que este los entendiera. Por ejemplo, hablaban en voz alta y lenta para que el comando se entendiese o reducían la complejidad del mismo, con ello aseguraban la ejecución precisa de la tarea. En otros casos cancelaban la acción y la empezaban nuevamente o eliminaban el evento para crearlo correctamente.

Ballati y otros [10], realizaron un estudio en donde hicieron uso de los asistentes de voz Google Assistant de Google, Siri de Apple y Cortana de Microsoft, todos ellos ejecutados sobre dispositivos móviles. En dicho estudio investigaron la medida en que personas diagnosticadas con disartria moderada la cual es un trastorno del habla que dificulta la articulación de palabras, podían ser entendidas por los asistentes antes mencionados. Para el estudio realizado, participaron 8 personas que presentan disartria, las cuales, grabaron 34 oraciones que fueron usadas sobre cada asistente. Con lo cual los investigadores buscaban conocer la comprensión de pregunta y consistencia de respuesta. La comprensión de la

pregunta fue analizada cuantitativa y cualitativamente, para la primera se revisó la transcripción de lo preguntado en pantalla y clasificado en 5 categorías. Los resultados muestran que Google transcribe con 64% aproximadamente de transcripción correcta mientras que Cortana y Siri lo hacen con un 40% y 15% respectivamente. Por otro lado, al evaluar la consistencia de la respuesta, es decir la parte cualitativa, los resultados muestran que Google y Siri responden de manera coherente entre un 50% y 60% mientras que Cortana de respuestas por defecto en un 76% aproximadamente.

Se encontró también estudios realizados por [11][12][13]. En ellos se han propuesto asistentes para personas con discapacidades visuales, los cuales hacen uso del procesamiento del lenguaje natural para reconocer las peticiones de los usuarios. La asistencia estuvo enfocada al reconocimiento del entorno para guiarlos, enviar y recibir mensajes, realizar búsquedas en internet entre otras. Las respuestas de las peticiones eran presentadas de forma auditiva.

Al finalizar la revisión bibliográfica se encontraron estudios en donde hacen uso de asistentes de voz ya sea en dispositivos móviles como en altavoces inteligentes. Dichos estudios han propuesto asistentes sobre dispositivos móviles mientras que otros han usado a los asistentes de voz comerciales sobre altavoces inteligentes y dispositivos móviles, encontrando limitaciones en la pronunciación y ejecución de comandos. Además de preferencias en cuanto se refiere a naturalidad de interacción.

Los resultados antes mencionados demuestran que, en la actualidad, no se han realizado investigaciones a fondo de asistentes de voz comerciales sobre altavoces inteligentes siendo usados por personas con diversidad funcional intelectual puesto que, la información encontrada evidencia la interacción de asistentes de voz a través de interfaces gráficas y en un único caso hacen uso de altavoces inteligentes empleados por personas con deterioro cognitivo. Dicho esto, la realización del futuro trabajo en donde se haga uso de altavoces inteligentes implica una interacción directa con interfaces de usuario por voz y para la cual aún no existen hallazgos referentes a esta alternativa de interacción con esta población.

3 Análisis técnico

Para la realización del estudio fue necesario un análisis de los altavoces a usar por lo cual se revisó las características de los principales asistentes de voz comerciales que funcionen sobre altavoces inteligentes.

3.1 Asistentes de voz

Los altavoces a utilizarse para el estudio son Alexa de Amazon, Siri de Apple y Google Assistant de Google. Existen otros asistentes comerciales que funcionan sobre altavoces inteligentes como Cortana y Bixby sin embargo no son usados en el estudio debido que a la fecha de realización del trabajo no se encuentran disponibles en español.

3.1.1 Siri

Es el asistente de voz propuesto por Apple y su lanzamiento fue en el año 2011. La forma de interactuar es a través de la voz para lo cual se necesita una frase de activación la cual es “Oye Siri”. El tipo de información devuelta por el asistente puede ser de manera gráfica o a través de respuestas por voz. En el caso de su altavoz inteligente toda la información es presentada a través de respuestas por voz. Dependiendo del dispositivo en el que se encuentre puede necesitar de un entrenamiento previo puesto que realiza una verificación de voz para su activación. La personalización de comandos es posible a través de la aplicación “Atajos” para lo cual es necesario escribir la frase a pronunciar para que ejecute el comando personalizado. Es el asistente por defecto de Apple por lo cual se encuentra disponible únicamente en dispositivos propios de la marca como ordenadores, dispositivos móviles entre otros. Algunos de los idiomas en los que se encuentra disponible este asistente son inglés, español, francés, chino, coreano, entre otros.

3.1.2 Google Assistant

Es el asistente de voz propuesto por Google, su año de lanzamiento fue en el 2016. La interacción con este asistente puede realizarse ingresando las solicitudes de manera textual o mediante comandos de voz para lo cual se utiliza una frase de activación “Ok, Google”, al realizar esta acción el asistente comienza a escuchar las peticiones a realizar. Las respuestas obtenidas por el asistente pueden darse de manera textual o respuestas por voz, esta última es utilizada en altavoces inteligentes que no poseen una pantalla. Este asistente se encuentra disponible en dispositivos con sistema operativo Android e IOS, para este último es necesario descargar la aplicación. Permite la personalización de comandos a través de la creación de rutinas para lo cual es necesario escribir el nuevo comando y si es necesario las variaciones del mismo. No requiere un entrenamiento previo para su utilización. Se encuentra disponible en idiomas como inglés, francés, español, entre otros.

3.1.3 Alexa

Es el asistente de voz propuesto por Amazon, el cual fue lanzado en el año 2014 sobre el altavoz inteligente Amazon Echo por lo cual, la interacción se realiza mediante comandos de voz. Actualmente este asistente se encuentra disponible en dispositivos con sistema operativo Android e IOS para la cual es necesario la instalación de la aplicación Alexa. Su frase de activación por defecto es “Alexa”, aunque la misma puede variar dependiendo del dispositivo en el cual se encuentra contenido. Posee la característica “skills”, la cual permite extender las habilidades del asistente a través de integración de funcionalidades de terceras partes. De igual manera permite personalizar comandos a través de la creación de rutinas para lo cual es necesario ingresar la frase exacta y la secuencia de tareas que realizara al

pronunciar el nuevo comando. Está disponible en idiomas como español, inglés, francés, italiano, entre otros.

3.2 Altavoces inteligentes

3.2.1 Homepod

Homepod⁵ es el altavoz inteligente propuesto por Apple, el cual contiene al asistente de voz Siri. Al activarse, presenta una luz en la parte superior del altavoz como se observa en la Figura 3.1, además, presenta una interfaz táctil, la cual permite aumentar o disminuir el volumen. Una alternativa para que el asistente comience a escuchar es mantener presionado en el centro de esta interfaz. La interacción que se mantiene con el asistente a través del altavoz es cifrada y enviada a los servidores de Apple de forma anónima sin asociar a la cuenta vinculada con el dispositivo por lo que no existe un historial de las peticiones realizadas al mismo.

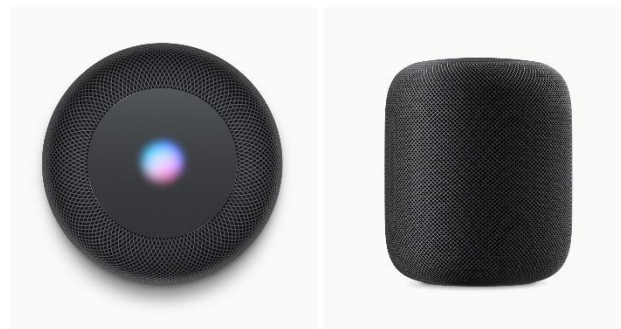


Figura 3.1. Altavoz Inteligente HomePod

Un resumen de especificaciones de este altavoz se encuentra en la Tabla 2, en donde especifica las características principales de este dispositivo.

Altavoz Homepod	
Características	Descripción
Frase de activación	Oye Siri
Control adicional	Tacto
Una pulsación	Reproducción/Pausa
Dos pulsaciones	Pista Siguiente
Tres pulsaciones	Pista anterior
Pulsación continua	Activación asistente
Signo "+"	Subir el volumen
Signo "-"	Bajar el volumen
Dimensión	172 mm alto 142 mm ancho
Cantidad de micrófonos	6
Activar/Desactivar micrófono	No disponible

Tabla 2. Características de HomePod

⁵ <https://www.apple.com/es/homepod/>

3.2.2 Google Home

El Google Home⁶, es uno de los modelos de altavoces propuestos por Google, el cual contiene a Google Assistant. Al utilizar la frase de activación muestra en la parte superior del dispositivo los colores de Google tal y como se presenta en la Figura 3.2. Presenta una interfaz que se activa con el tacto la cual permite subir o bajar el volumen al deslizar los dedos en sentido horario o anti horario respectivamente. Además, posee un botón en la parte posterior el cual permite activar o desactivar el micrófono del dispositivo[14][14]. Toda la interacción que se mantiene con el asistente se almacena en el historial de la cuenta con la cual ha sido vinculada el altavoz, lo que permite revisar las peticiones realizadas al mismo.



Figura 3.2. Altavoz Inteligente Google Home

Un resumen de especificaciones de este altavoz se encuentra en la Tabla 3, en donde especifica las características principales de este dispositivo.

Altavoz Google Home	
Características	Descripción
Frase de activación	Ok Google
Control adicional	Tacto y botón
Una pulsación	Reproducción/Pausa/Detener alarmas o temporizadores
Pulsación continua	Activación asistente
Giro sentido horario	Subir el volumen
Giro sentido anti horario	Bajar el volumen
Dimensión	142.8 mm alto 96.4 mm ancho
Cantidad de micrófonos	2
Activar/Desactivar micrófono	Botón silencio

Tabla 3. Características de Google Home

⁶ https://store.google.com/es/product/google_home_specs

3.2.3 Amazon Echo

El modelo Amazon Echo⁷ contiene al asistente de voz Alexa. Al utilizar la frase de activación el dispositivo muestra en la parte superior un anillo de luz en la parte superior del altavoz tal y como se presenta en la Figura 3.3, la cual apunta en dirección desde donde el usuario invoca al asistente.. Las variables de frase de activación para este dispositivo son “Amazon” o “Echo”. Al igual que en Google Assistant, toda la interacción que se mantiene con Alexa, se almacena en el historial de la cuenta con la cual ha sido vinculado el altavoz, lo que permite revisar las peticiones realizadas al mismo.



Figura 3.3. Altavoz Inteligente Amazon Echo

Un resumen de especificaciones de este altavoz se encuentra en la Tabla 4, en donde especifica las características principales de este dispositivo.

Altavoz Amazon Echo	
Características	Descripción
Frase de activación	Alexa
Control adicional	Botones
Botón "."	Reproducción/Pausa Activación asistente
Botón "+"	Subir el volumen
Botón "-"	Bajar el volumen
Dimensión	148 mm alto 88 mm ancho
Cantidad de micrófonos	7
Activar/Desactivar micrófono	Botón silencio

Tabla 4. Características de Amazon Echo.

3.3 Funcionalidades comunes

Puesto que el estudio utilizó los altavoces anteriormente mencionados, fue necesario revisar los sitios web oficiales de las empresas propietarias de los asistentes de voz con lo cual, se generó una lista de funcionalidades que fueran capaces de ser realizadas por los asistentes sobre altavoces inteligentes. Adicional a ello, se revisó en foros y webs que mostraran información de los altavoces de Google, Apple y Amazon.

⁷ <https://www.aboutamazon.de/innovationen/amazon-echo>

A continuación, se realizaron pruebas en el laboratorio de inteligencia ambiental AMILAB de la Universidad Autónoma de Madrid, lugar en el cual, se pudo verificar el correcto funcionamiento de las funcionalidades comunes encontradas entre los asistentes de voz sobre altavoces inteligentes. La lista completa se muestra a continuación:

- **Configuración de alarmas:** Permite la configuración de alarmas, es necesario indicar la hora y de manera opcional el día.
- **Configuración de recordatorios:** A diferencia de las alarmas, los recordatorios varían su manifestación puesto que, al indicarlos desde el altavoz, este los muestra en el móvil desde donde se encuentre la cuenta vinculada al altavoz inteligente en el caso de Amazon y Google.
- **Configuración de temporizadores:** Esta funcionalidad se asemeja a una alarma con la diferencia de que se indica el rango de tiempo hasta su activación.
- **Gestión de eventos en el calendario:** Permite agendar eventos en el calendario, es necesario indicar la fecha, hora y nombre del evento. Cuando se realiza la petición de manera incompleta, el asistente solicita la información faltante.
- **Obtención de fechas:** Permite obtener la fecha actual, así como también sobre acontecimientos de carácter general.
- **Obtención del tiempo:** Esta funcionalidad permite realizar peticiones sobre el tiempo en específico de un lugar, caso contrario responde el clima con la ubicación de la configuración
- **Manejo de calculadora:** Permite realizar operaciones sencillas a través de la voz.
- **Obtención de información:** Esta funcionalidad permite realizar preguntas de interés del usuario.
- **Obtención de direcciones:** Esta funcionalidad permite consultar direcciones y hace uso de una configuración inicial en donde se indica la ubicación del altavoz.
- **Reproducción de música:** Permite reproducir música dando como parámetro el nombre del artista o canción. Es necesario tener un servicio de música en Streaming.
- **Traductor:** Permite realizar traducciones en donde es necesario indicar a que idioma y que frase debe traducir.
- **Interacción con otros dispositivos:** La funcionalidad de interacción con otros dispositivos como luces o enchufes depende del fabricante puesto que hay dispositivos específicos para cada altavoz.
- **Preguntas interactivas:** La funcionalidad de preguntas interactivas se refiere a peticiones al altavoz de contar chistes, cuentos, color favorito, etc.
- **Personalización de comandos:** Permite la personalización de comandos con lo cual permite ejecutar tareas al decir frases específicas. Esta función varía entre los asistentes, con Siri es llamado Atajos, mientras que con Alexa y Google Assistant es llamado Rutinas.

Existen otras funcionalidades como enviar y recibir mensajes/emails, conversación continua, pero estas no han sido incluidas debido a que, a la fecha de redacción de este documento, estas funcionalidades están disponibles en otros idiomas o no todos los altavoces pueden hacerlas. Por ello, se tomaron en cuenta solo aquellas que pudieran ser ejecutadas con los 3 altavoces.

3.4 Dispositivos adicionales

Para la realización del estudio se harán uso de dispositivos que permitan probar las funcionalidades de los altavoces con otros dispositivos para lo cual se realiza una descripción de los mismos.

3.4.1 Enchufe Koogeek P1

El dispositivo mostrado en la Figura 3.4, es un enchufe inteligente, el cual posee una conexión inalámbrica. Permite encenderlo y apagarlo a través de los asistentes de voz de Apple, Google y Amazon. Posee también un botón el cual permite encenderlo o apagarlo manualmente. La configuración de este dispositivo es realizada a través de la aplicación propia de la marca. Luego de ello es necesario sincronizarlo con cada asistente. La sincronización se realiza mediante las aplicaciones HomeKit de Apple, Google Home de Google y Amazon Alexa de Amazon. La configuración anterior permite personalizar el nombre del enchufe para ser usado con los comandos de voz.



Figura 3.4. Enchufe Koogeek P1

3.4.2 Ventilador

Se hará uso de un ventilador como el mostrado en la Figura 3.5, el cual, va conectado con el enchufe anterior lo cual permitiría encenderlo y apagarlo a través de comandos de voz. Se hará uso de lazos de colores atados en el mismo, con lo cual les resultase amigable a los participantes del estudio.



Figura 3.5. Ventilador a usarse con enchufe.

3.4.3 Phillips Hue Bridge

El dispositivo mostrado en la Figura 3.6, permite la comunicación entre los asistentes de voz y dispositivos de la marca Phillips. La conexión de este dispositivo es a través de un cable de red conectado directamente al enrutador en donde se encuentran los demás dispositivos enlazados al mismo. De igual manera que como en el enchufe, se requiere una configuración inicial la cual se realiza con la aplicación propia de la marca para luego sincronizarlo con las aplicaciones de los asistentes de voz.



Figura 3.6. Dispositivo Phillips Hue Bridge

3.4.4 Phillips Hue Go

El dispositivo Phillips Hue Go mostrado en la Figura 3.7, es una luz inteligente la cual posee una conexión inalámbrica. Permite ser encendida o apagada de manera manual o a través de comandos de voz. Este dispositivo es compatible con los asistentes de Apple, Amazon y Google, con lo cual, luego de ser enlazada con el dispositivo mostrado en la Figura 3.6, se enlaza automáticamente con las aplicaciones de los asistentes en donde el dispositivo antes mencionado se haya configurado. Posee una batería integrada la cual permite mantenerla sin conexión eléctrica. Para la realización de las sesiones, se mantendrá conectada.



Figura 3.7. Lámpara Phillips Go

4 Metodología

La metodología a aplicar en el estudio es *Coaching Method*[15]. Este método de evaluación es seleccionado debido a que, los participantes del estudio son usuarios inexpertos con lo cual, en todo momento serán guiados en la realización de las tareas con los diferentes altavoces. De igual manera el evaluador responderá preguntas que surjan por parte de los participantes. Para realizar el estudio, es necesario la definición de tareas a realizar. Además de formar grupos tomando en cuenta las capacidades de cada participante por lo cual se trabajó conjuntamente con el personal docente y administrativo de ALENTA, el cual es un centro de educación especial que brinda apoyo en las distintas etapas de desarrollo de personas con diversidad funcional intelectual, con lo cual se generaron tareas de acuerdo a las capacidades y conocimiento de los participantes. Finalmente se definieron variables que permitirían evaluar tanto la pronunciación de la frase de activación, así como también la formulación, pronunciación y respuesta obtenida por el asistente.

De igual manera se realizará el estudio con participantes con un perfil tecnológico con lo cual realizarán las mismas tareas que los participantes del primer grupo, sin embargo, se entregará en papel las tareas a realizar por asistente. Al igual que con los participantes de ALENTA, el evaluador responderá cualquier inquietud que se presente durante las sesiones.

Por otro lado, se realizarán entrevistas para obtener una retroalimentación sobre las limitaciones y ventajas de utilizar los altavoces. Las entrevistas con los participantes de ALENTA serán realizadas a los educadores puesto que son ellos quienes conocen a sus alumnos y pueden brindar una visión general de acontecimientos que noten durante las sesiones. Los participantes con perfil tecnológico realizarán la entrevista al finalizar la sesión de manera individual puesto que pueden aportar mayor información debido a su conocimiento en el área.

4.1 Procedimiento

El procedimiento a seguir para realizar el estudio es Latin Square[16]. Esta metodología es propuesta puesto que permite contrastar las diferencias entre los diferentes altavoces realizando las mismas actividades y evita la generación de sesgos por la presencia de varios grupos de participantes. En el estudio se utilizan 3 altavoces diferentes y se definirán 3 grupos de tareas. Con lo cual será necesario la segmentación de los participantes del estudio en 3 grupos. En la Figura 4.1, se presenta la configuración del procedimiento. Como se puede observar cada grupo de participantes realizará un grupo de tareas con un altavoz diferente. Con ello se asegura que cada participante utilice los 3 altavoces realizando un grupo de actividades en cada uno.

4.2 Tareas

La definición de tareas se realizó en conjunto con educadores de ALENTA, con lo cual en reuniones se presentó las funcionalidades en común que pueden realizar los 3 altavoces. Los educadores consideraron las funcionalidades más relevantes para los participantes puesto que indicaron que no tendría sentido gestionar el calendario o preguntar por direcciones. Al finalizar las revisiones se establecieron 3 grupos de tareas. Cada una de ellas realizan tareas específicas y se describen en los puntos siguientes.

	G. Tareas 1	G. Tareas 2	G. Tareas 3
G. Participantes 1	Homepod	Amazon echo	Google Home
G. Participantes 2	Amazon echo	Google Home	Homepod
G. Participantes 3	Google Home	Homepod	Amazon echo

Figura 4.1. Procedimiento usando Latin Square

4.2.1 Grupo de tareas 1

En este primer grupo de tareas está enfocado en la interacción con otros dispositivos puesto que una de las características que presentan los asistentes sobre los altavoces es la de controlar otros dispositivos como enchufes, luces, termostatos, persianas entre otros. Con lo cual se definen las siguientes tareas:

- Tarea 1: Encender una luz
- Tarea 2: Poner un temporizador de 30 segundos
- Tarea 3: Encender un ventilador
- Tarea 4: Detener el temporizador
- Tarea 5: Apagar la luz
- Tarea 6: Apagar el ventilador

En este grupo de tareas, los educadores manifestaron que configurar el temporizador no tendría sentido para ciertos participantes por lo que esta tarea será usada de acuerdo a las indicaciones del educador presente en el momento de la sesión.

4.2.2 Grupo de tareas 2

El segundo grupo de tareas corresponde a obtener información específica, con lo cual se proponen 4 tareas a realizar las cuales hagan uso de diferentes funcionalidades que son comunes en los asistentes de voz a usarse en el estudio. Se considera usar la pregunta referente a una festividad debido a que los participantes tendrán vacaciones en dicha fecha.

- Tarea 1: Preguntar la hora
- Tarea 2: Preguntar el tiempo
- Tarea 3: Realizar una suma
- Tarea 4: Preguntar sobre ¿Cuánto falta para semana santa?

De igual manera que en el grupo anterior, los educadores manifestaron que algunos participantes no realizarían la suma o preguntarían sobre fechas específicas.

4.2.3 Grupo de tareas 3

El tercer grupo de tareas está pensado en que el usuario pueda realizar peticiones de temas de interés personal. En el caso de que no se les ocurra que preguntar en ese momento, se le sugerirán preguntas de acuerdo a su edad o conforme lo que sugiera el educador que se encuentre presente en la sesión. Como mínimo se pedirá a los usuarios que realicen 3 peticiones, sin embargo, son libres de realizar las que deseen. Algunas sugerencias son las listadas a continuación:

- Pedir música
- Contar chistes
- Contar cuentos
- Preguntar sobre deportes

Adicional a ello, una pregunta propuesta por los educadores fue la de preguntar sobre el menú del día debido a que es algo que preguntan con regularidad. Fue necesario añadir un comando adicional en los asistentes para que pudieran contestar a la pregunta, sin embargo, es necesario describir específicamente como realizar la petición, para lo cual, las variaciones de realizar esta pregunta fueron planteadas por los educadores y son las siguientes:

- Que hay de comer
- Que hay de menú
- Hoy que comer
- Que vamos a comer hoy
- Que hay hoy de comer
- Que comemos hoy

4.3 Sesiones

Para realizar el estudio se propone realizar 2 tipos de sesiones con los participantes de ALENTA, las cuales permitan realizar una evaluación exitosa. Por otro lado, se realizará una sola sesión a los participantes con perfil tecnológico puesto que, como se explicará a continuación, la primera sesión está enfocada en presentar los altavoces a usuarios no expertos.

4.3.1 Sesión de entrenamiento

Esta sesión se propone con el fin de presentar los altavoces inteligentes a los participantes puesto que muchos de ellos desconocen del funcionamiento de estos dispositivos y la forma de interactuar con los mismos. Esta primera sesión es realizada de manera grupal con lo cual se busca que los participantes se sientan cómodos al encontrarse con compañeros y educadores del centro. Por otro lado, se busca que los participantes se familiaricen con el grupo de investigación puesto que, la realización de las sesiones de estudio será efectuadas de manera individual con la presencia del educador y los evaluadores.

4.3.2 Sesión de estudio

Con el fin de evitar distracciones o confusiones por parte de los participantes debido a la frase de activación o forma del altavoz se propone realizar de manera individual el grupo de tareas del altavoz sujeto a pruebas en ese momento con el pictograma correspondiente. Al finalizar el primer grupo de tareas respectivo se procederá a guardar el primer altavoz y los materiales adicionales que hayan sido usados y se procederá con el segundo altavoz y materiales correspondientes. Se procederá de igual manera con el tercer altavoz.

Por otro lado, se dará un número de 3 intentos tanto para la activación del asistente como para la ejecución de la tarea. En caso de pasar estos intentos el evaluador activará al altavoz o ejecuta la tarea según sea el caso. Estas acciones tienen 2 propósitos. El primero es dado por si los participantes presentan problemas al utilizar la frase de activación, pero pueden ejecutar la tarea o pueden pronunciar la frase de activación, pero presentan dificultad al ejecutar la tarea propuesta. El fin del segundo propósito es el de que el participante tenga un refuerzo positivo y se eviten frustraciones por no poder realizar las tareas que se le pidan y con ello poder continuar con el estudio. Al finalizar la sesión con cada participante se revisará la lista de participantes para las sesiones del día y se procede a ubicar el altavoz correspondiente con el siguiente participante.

4.3.3 Duración de las sesiones

Para determinar el tiempo requerido se mantuvieron reuniones con los educadores de ALENTA los cuales indicaron que las sesiones durarían alrededor de 30 minutos, debido a que consideraron que era un tiempo prudente y que no era posible extenderlo debido a la frustración o pérdida de atención que tuviesen durante las mismas. Los educadores informaron que para ciertos participantes era conveniente realizar las sesiones en varios días puesto que su tolerancia a la frustración era baja con lo que en el peor de los casos abandonarían el experimento al encontrarse bajo mucha tensión.

La duración de las sesiones a realizar por los participantes será de un tiempo estimado de 15 minutos. Este tiempo ha sido determinado puesto que se realizaron pruebas en el laboratorio de inteligencia ambiental AMILAB, haciendo uso de las tareas propuestas en el punto 4.2.

4.3.4 Protocolo de sesiones

Para la realización de las sesiones se establece un protocolo el cual era verificado al inicio de cada sesión y consiste de los siguientes pasos:

Altavoces encendidos: Se verifica que los 3 altavoces estén encendidos. Esta comprobación se realizaba para evitar problemas durante las sesiones.

Conexión a internet: De igual manera, se verifico la conexión de internet de cada altavoz realizando preguntas en las que era necesario realizar una consulta y devolverla al evaluador. Preguntas como “*Que tiempo hace*” o “*Reproduce música*” fueron usadas para realizar esta comprobación.

Revisión de videgrabadora: Se verifico en el inicio de cada sesión que hubiese suficiente capacidad de almacenamiento, fuente de energía conectada, nivel de batería suficiente, volumen adecuado e inicio de grabación con lo cual se aseguró de no perder ningún detalle de las sesiones

Comprobación de dispositivos adicionales: Se comprobará que el enchufe y la luz funcionen correctamente puesto que, en varias ocasiones los dispositivos antes mencionados dejan de funcionar de manera inesperada.

4.4 Análisis de Resultados

Con las grabaciones realizadas durante las sesiones se realizará un análisis sobre las frases de activación usadas por cada asistente así también como las tareas a realizar y las respuestas obtenidas.

4.4.1 Frases de activación

El análisis propuesto, tomara en cuenta el número de intentos que el participante necesite para activar los altavoces puesto que se busca encontrar las limitaciones o ventajas que puedan existir sobre las frases de activación de los asistentes de voz comerciales.

4.4.2 Tareas realizadas

Este análisis será realizado tomando en cuenta la orden dada por el participante y comparada con la respuesta esperada por el altavoz. Para lo cual será necesario tomar en cuenta en el análisis de las grabaciones, 2 variables las cuales será formulación de tarea y respuesta obtenida de la tarea.

4.4.2.1 Formulación de la tarea

- Petición Correcta: Cuando luego de explicarle al usuario que tarea debe realizar, el mismo lo ejecuta de forma correcta y con las palabras esperadas y/o adecuadas.
- Petición Incompleta: De igual manera que la anterior, sin embargo, al pronunciar la tarea lo hace con palabras incompletas o en el caso de omisión de palabras como artículos, etc.
- Petición incorrecta: Cuando se pronuncian otras palabras que no tienen nada que ver a lo solicitado.

4.4.2.2 Respuesta de la tarea

- Correcta: Cuando la respuesta obtenida por el altavoz es la acertada en relación a lo preguntado.
- Defecto: Respuesta generada por defecto por el altavoz cuando no entiende una petición, se incluye además la omisión de respuesta puesto que en ciertos casos el altavoz no devuelve una respuesta por voz.
- Incorrecta: Cuando la respuesta obtenida por el altavoz no es la esperada.

5 Evaluación

5.1 Participantes

Para la realización del estudio, se contó con dos grupos de participantes. El primer grupo estuvo conformado por estudiantes y residentes del centro de educación especial ALENTA. El segundo grupo estuvo conformado por estudiantes y profesores de la Escuela Politécnica Superior de la Universidad Autónoma de Madrid los cuales tienen un perfil tecnológico. Quienes participaron en el estudio, firmaron un consentimiento de participación y derecho de imagen, en el caso de quienes están considerados dentro del entorno de diversidad funcional intelectual, fue necesario el consentimiento por parte de sus representantes.

5.1.1 Participantes ALENTA

Un total de 26 personas de edades entre 9 y 50 años participaron en el estudio de los cuales, 7 eran mujeres y 19 hombres. Además 10 de los participantes son de residencia y los 16 restantes son alumnos del centro. Por otro lado, se obtuvieron las características de cada participante las cuales se muestran en la Tabla 6. Esta tabla contiene información demográfica y etnográfica usadas en el centro tales como grado de discapacidad intelectual, nivel de pronunciación, nivel comunicación nivel de atención y nivel de comprensión. Dichos niveles están representados de forma numérica los cuales van entre el nivel más severo correspondiente a 1 y el más leve correspondiente a 3. Salvo el nivel de pronunciación que va en un rango de 1 a 5. Incluye información adicional como la edad y el género y finalmente el nivel en el cual se encuentran dentro de ALENTA. En algunos casos, se incluye información adicional sobre el diagnóstico.

5.1.2 Participantes con perfil tecnológico

Por otro lado, se contó con 21 participantes de la Escuela Politécnica Superior de los cuales 11 son mujeres y 10 hombres. Los participantes de este grupo provienen de diferentes países y hablan español. Con lo cual pueden ayudar a corroborar posibles limitaciones con los participantes del grupo con diversidad funcional intelectual, además de obtener sus opiniones sobre los altavoces inteligentes y los respectivos asistentes que contienen.

5.2 Materiales

5.2.1 Altavoces inteligentes

Como se revisó en el capítulo de análisis técnico sobre altavoces inteligentes, en la Tabla 5, se muestran el modelo de altavoz, el asistente que contiene y su frase de activación.

Altavoces Inteligentes		
Modelo Altavoz	Asistente contenido	Frase de activación
HomePod	Siri	"Oye Siri"
Google Home	Google Assistant	"Ok Google"
Amazon Echo	Alexa	"Alexa"

Tabla 5. Altavoces usados en la evaluación

Grupo	ID Participante	Género	Edad	Nivel	Diagnostico	Grado	Pronunciación	Comunicación	Atención	Comprensión
1	1	H	46	3	N/D	76%	3	1	2	2
	2	M	47	3	N/D	65%	5	3	3	3
	3	H	38	3	N/D	77%	5	3	3	3
	4	H	12	2	DI	Moderada	3	1	3	2
	5	M	15	1	DI	Moderada	4	2	2	2
	6	M	17	2	DI	Moderada	3	2	2	2
	7	H	N/D	2	DI	Moderada	5	3	3	3
	8	H	15	1	DI	Moderada	3	2	2	2
2	9	H	45	3	N/D	65%	2	1	2	2
	10	H	30	3	N/D	75%	5	1	2	3
	11	H	43	3	N/D	67%	5	3	3	2
	12	H	9	2	DI	Moderada	3	2	3	2
	13	H	15	1	TEA	N/D	5	2	1	1
	14	M	19	2	DI	Moderada	3	3	2	3
	15	H	12	1	DI	Moderada	5	3	3	3
	16	H	16	1	TEA	N/D	4	2	2	2
	17	H	13	1	TEA	N/D	5	2	1	1
3	18	H	50	3	N/D	65%	5	2	2	2
	19	M	35	3	N/D	65%	5	2	2	2
	20	H	32	3	N/D	81%	5	3	3	3
	21	M	41	3	N/D	68%	4	3	3	3
	22	H	14	1	N/D	N/D	N/D	N/D	N/D	N/D
	23	H	15	1	TEA	N/D	4	1	1	1
	24	H	21	2	DI	Moderada	5	3	3	3
	25	H	14	1	TEA	N/D	5	3	3	3
	26	M	17	1	TEA	N/D	5	2	1	1

Tabla 6: Información de participantes de ALENTA

5.2.2 Configuraciones altavoces inteligentes

Para la realización de las pruebas, se mantuvo la configuración estándar de cada dispositivo. El volumen usado en las pruebas fue un nivel medio. Además, aunque Amazon Echo permite seleccionar entre varias frases de activación, se dejó esta configuración por defecto, siendo Alexa la frase a utilizar en la evaluación. Por otro lado, los asistentes permiten configurar el género de la voz, pero de igual manera se mantiene las voces que vienen por defecto.

5.2.3 Videgrabadora

Se utilizó además una videgrabadora provista de una memoria de 128 gigabytes y 2 memorias adicionales para evitar problemas de falta de almacenamiento.

5.2.4 Configuración del entorno de evaluación

Puesto que los altavoces se comunican con los dispositivos si se encuentran en la misma red, fue necesario la utilización de un dispositivo de enrutamiento para mantener a todos los dispositivos dentro de la misma y que funcionaran sin problemas. Para cada sesión se realizaba las conexiones correspondientes y se verificaba que todos los dispositivos funcionaran correctamente.

5.2.5 Pictogramas

Se hizo uso de pictogramas para realizar las sesiones de prueba. En la Figura 5.1, se muestran los pictogramas usados para llamar a los asistentes los cuales fueron creados por educadores de ALENTA. En el caso de los altavoces de Apple y Google se usan figuras que corresponden a partes de la frase de activación, las cuales son familiares para los participantes. En el caso de Alexa no hace falta puesto que es una sola palabra para activar el altavoz.



Figura 5.1. Pictogramas de altavoces inteligentes

De igual manera para la realización de las tareas, los educadores crearon pictogramas los cuales se muestran en el Anexo A. Estos pictogramas brindaron apoyo para la realización de ciertas tareas debido a que los participantes hacen uso de este tipo de material. En otros casos era necesario únicamente explicarles en qué consistía la tarea para que la realizaran.

5.3 Sesiones ALENTA

5.3.1 Sesión Entrenamiento

Previamente a la realización de las sesiones de prueba, se mantuvo la sesión de entrenamiento grupal con los participantes y su educador entre los meses de marzo y abril, para que pudieran conocer la manera de interactuar con los altavoces puesto que son algo nuevo para ellos y el concepto que tienen a mantener una conversación va ligado a referirse a una persona. Estas sesiones duraron entre 30 y 40 minutos.

Durante estas sesiones grupales, se preguntó a cada participante cuál era su nombre. Luego de ello se presentaron a los altavoces por el nombre del asistente que contiene en conjunto con el pictograma para que se familiaricen con ellos. A continuación, se explicó la forma de usar la frase de activación para llamar a los asistentes y se explicó a los participantes, las señales visuales de cuando estos han comenzado a escucharlos. Se realizaron peticiones de prueba como “*¿Qué día es hoy?*” o “*Reproduce música*” para que entendiesen que es lo que podían hacer los asistentes por ellos. Finalizada estas pruebas se permitió que cada uno de ellos realizaran las peticiones que les interesara con cada uno de los asistentes para que probaran la forma de llamarlos. En algunos casos en donde los participantes no intervenían en la sesión se dio sugerencias de que podían preguntar. Los participantes de residencia realizaron preguntas sobre deportes que les gustaban o su música favorita.

Por otro lado, en las sesiones grupales realizadas por los participantes del centro de educación, se utilizó un gorro para tomar turnos de participación puesto que, en un principio, todos querían hacer uso de los altavoces. Varias de las solicitudes de los participantes eran sobre canciones de su gusto con lo cual cantaban y en el caso de los participantes más jóvenes bailaban al escuchar lo que habían pedido. Los participantes realizaban las peticiones con gran cordialidad para lo cual los educadores indicaron que algo que les enseñan dentro de ALENTA es la amabilidad al pedir algo.

Al finalizar estas sesiones se preguntó a los participantes si les había gustado usar los altavoces, a lo cual indicaron que habían disfrutado mucho conocerlos, e incluso uno de ellos preguntó el valor de los altavoces puesto que le pediría que su padre le comprara uno. Finalmente se indicó que podrían volver a utilizar los altavoces en las sesiones que se habían programado para días siguientes.

5.3.2 Sesiones de Prueba

Las pruebas se realizaron entre los meses de abril y mayo del 2019. Las sesiones fueron llevadas a cabo en días de clases con lo cual, el personal administrativo de ALENTA, planteo un cronograma para la realización de las mismas el cual se muestra en la Tabla 7.

Planificación participantes ALENTA					
Mes	Abril				Mayo
Día	Jueves 4	Lunes 8	Miércoles 10	Jueves 11	Martes 7
10:30	P4	P6	P17		
11:00	P12	P13	P16		P11
11:30	P24	P26	P23	P3	P19
12:00	P7	P5	P15	P20	
12:30	P14	P10	P22		
13:00			P9		
14:00			P2		
14:30	P17		P21		
15:00	P8		P3		
15:30	P25				

Tabla 7: Cronograma de sesiones en ALENTA

5.3.2.1 Inicio de sesión

Al iniciar las sesiones con los participantes, se les pregunto si recordaban el nombre del asistente que era objeto de pruebas ese momento, de igual manera se mostraba el pictograma perteneciente a cada asistente, con lo cual los participantes revisaban como interactuar con el altavoz. Se realizaron pruebas, pronunciando la frase de activación con lo cual, se explicó nuevamente las señales visuales que los altavoces muestran cuando son llamados.

5.3.2.2 Durante la sesión

Durante las sesiones, los participantes ejecutaban el comando de activación en donde se pudo evidenciar problemas con las frases de activación “Oye Siri” y “Ok Google” debido a que, usan 2 palabras y varios participantes solo usaban la que se refería al nombre o en el caso de Alexa, presentaban problemas de vocalización con lo cual, como se explicó en la metodología, el grupo de investigación ejecuto la frase de activación y el participante la tarea a realizar. En algunas sesiones los participantes tocaban los botones de los asistentes por lo que en el caso de Siri, comenzaba a reproducir música o subían o bajaban el volumen.

Primer grupo de tareas

Al comenzar este grupo de tareas se explicó a los participantes que con los asistentes podían controlar la luz o el ventilador. Se indicaba que debían pedir al asistente que encendiera y/o apagara la luz o el ventilador. En algunas ocasiones, los participantes realizaban las peticiones de encendido o apagado diciendo únicamente, “enciéndelo” o “apágalo” mientras apuntaban o fijaban la mirada en estos dispositivos para lo cual se les explicaba que debían aclarar que dispositivo querían encender o apagar.

En la tarea del temporizador se pedía a los participantes que pidieran al asistente que ponga un temporizador de 30 segundos. En algunos casos, pronunciaban la mitad de la petición para lo cual el asistente pedía que se confirme el tiempo del temporizador. Cuando el temporizador comenzaba a sonar, se indicaba que debían pedirle al asistente que lo detenga. Algunas de las frases que usaron para realizar esta tarea fueron apágalo, detente, para.

En ciertas sesiones existieron problemas por parte de los dispositivos para realizar las tareas de encendido y apagado debido a que, al dar las ordenes de encender y/o apagar la luz o el ventilador, los dispositivos retornaban mensajes por problemas de comunicación con los otros dispositivos con lo cual se realizó la tarea manualmente para continuar con las sesiones.

Segundo grupo de tareas

De igual manera al comenzar con este grupo de tareas se explicaba a los participantes que podían preguntarles a los asistentes sobre el tiempo. En los casos que era necesario se utilizó el pictograma mostrado en la Figura 8.5. Las respuestas obtenidas al preguntar sobre el tiempo daban demasiada información pese a que se preguntó a los participantes cual iba a ser el tiempo luego de escuchar la respuesta por parte del asistente, muchos de ellos se guiaban por alguna frase escuchada o por lo que habían visto afuera

Luego de esta tarea se les pedía que preguntasen la hora y al obtener la respuesta se verificaba si habían escuchado. Algunos participantes que traían reloj verificaban la respuesta obtenida por el asistente. Al realizar la tercera tarea, se explicó a los participantes que los asistentes sirven como calculadora con lo cual se pidió que realicen la operación. En esta tarea existieron problemas al ejecutar la petición debido a que, como se especifica en la Tabla 6, los participantes presentan barreras de vocalización por lo cual, tuvieron que repetir la tarea para conseguir el resultado.

Finalmente, en la última tarea de este grupo se les pregunto si sabían cuando tenían vacaciones para lo cual se les explico que preguntaran a los asistentes cuanto falta para Semana Santa. La respuesta obtenida varía entre los asistentes puesto que en el caso de Alexa y Siri, daban una respuesta detallada de la pregunta referente a la Semana Santa, por otro lado, el asistente de Google retornaba un número como respuesta por lo que era difícil de interpretar para los participantes.

Tercer grupo de tareas

En la sección de conversación abierta, se indicó a los participantes que podían preguntar sobre temas de su interés. Varios participantes realizaron preguntas personales refiriéndose a sus amigos, familiares o sobre información personal, con lo cual los asistentes respondían que no han entendido, o que no saben la respuesta. En este caso se explicó a los participantes que los asistentes no sabían esa información. En otros casos preguntaban sobre cuando jugaban sus equipos de futbol favoritos o pedían las canciones que más les gustaban. Por otro lado, los participantes 15 y 22 prepararon una lista de preguntas a realizar con los asistentes las cuales eran referentes a lugares que conocían o lugares a donde iban a ir de vacaciones con lo cual, aunque la formulación de las preguntas era la correcta, los asistentes no podían responder de manera adecuada y devolvían respuestas por defecto. Al usar el pictograma mostrado en la Figura 8.6, el cual pide que se cuente un cuento, los asistentes Alexa y Google responden contando cuentos conocidos para los participantes del centro de educación. Por el contrario, Siri daba respuestas como *“Seguro que ya te sabes todas las historias”* con lo cual se evidencio incertidumbre en los participantes por la respuesta obtenida. Mientras que, al usar el pictograma mostrado en la Figura 8.7, en la cual se pide al asistente que cante una canción, se evidencio la diferencia de respuesta puesto que, en el caso de Alexa cantaba canciones conocidos para niños mientras que Google Assistant cantaba canciones propias sobre el mismo. En el caso de Siri, respondía que cantar era muy difícil y aunque se considera como una respuesta válida, el descontento de los participantes era evidente. Al ejecutar comandos personalizados, en este caso el comando referente a preguntar sobre la comida del día, en los casos en el que la formulación o pronunciación de la misma no era exacta, los asistentes sugerían lugares cercanos de donde podrían comer lo cual causo desconcierto a los usuarios.

Problemas durante la sesión.

El educador de los usuarios 8, 16, 17 y 25 manifestó durante las sesiones que, los participantes necesitan un tiempo más largo para pronunciar la tarea, además de que tienen la voz baja. Recalco que, pese a que el asistente de voz Siri, capta con mayor facilidad a sus alumnos, las respuestas que da son poco amigables lo cual causa frustración o desinterés en realizar el estudio. En el caso del usuario 8 y 17 se realizaron tareas que constituyan un refuerzo positivo puesto que en las sesiones que realizaron tuvieron bastante dificultad al interactuar con los dispositivos.

El usuario 17 realizó 2 sesiones en días diferentes debido a que su educador indicó que no soportaría realizarla en 30 minutos. En la primera sesión se realizaron el grupo de tareas de interacción con dispositivos y el grupo de tareas de conversación abierta. La segunda sesión se realizó con Google Assistant para culminar con la sesión del usuario, sin embargo, se presentaron varios inconvenientes en las mismas. En la primera existieron problemas de conexión con el ventilador usado puesto que al dar la orden de que se pagara, el asistente indicó que el dispositivo no responde. En la segunda sesión el usuario respondía a las peticiones que debía preguntarle al asistente. Uno de los ejemplos que se dieron fue cuando se le pidió que preguntara el tiempo, pese a que tenía los pictogramas, el respondía que, hacia sol, o al pedirle que preguntase la hora el participante respondía 10:30. Para evitar problemas de frustración, el educador pidió realizar tareas de contar cuentos para que el participante tuviera un refuerzo positivo.

La presencia de los educadores durante las sesiones fue de gran ayuda debido a que, en las mismas, notaban cuando sus alumnos sentían vergüenza o frustración al realizar las tareas. Señales difíciles de interpretar para el grupo de investigación, con lo cual, los educadores les indicaban que se tranquilicen, que estaban en confianza y que lo hicieran tranquilos. Algunos casos en específico, en donde el educador asignado intervino se dio con los participantes 4, 5, 10, 12, 17, 18, 23. En otros casos como con el usuario 8, era evidente la frustración causada al no poder activar ni ejecutar las tareas a realizar por lo que en este caso obtuvo ayuda del grupo de investigación para que tuviera un refuerzo positivo de la sesión.

5.3.3 Entrevista realizada a educadores de ALENTA

Se llevó a cabo una entrevista con los educadores, con lo cual se pretendió obtener una visión general sobre los problemas que hayan enfrentado sus alumnos durante las sesiones. Al preguntarles sobre con cual o cuales frases de activación sus alumnos habían podido activarlas con mayor facilidad, los educadores respondieron que Alexa y Oye Siri eran más sencillos de pronunciar. De igual manera se pregunta cual frase había presentado mayores problemas con sus alumnos a lo que respondieron en mayor medida Google y Siri.

Esta contradicción entre sus respuestas se da debido a que sus alumnos presentan diversas limitaciones en la pronunciación con lo cual, a su percepción, los fonemas empleados y el uso de 2 palabras para activar a los asistentes les presenta un desafío en su utilización.

Por otro lado, al preguntarles sobre las preguntas realizadas por parte de los alumnos a los asistentes, los educadores comentaron que, en algunos casos, las respuestas obtenidas contienen demasiada información y un lenguaje muy técnico con lo cual la comprensión de las respuestas es mínima.

Al preguntarles sobre la percepción de activación de los altavoces, respondieron que Alexa ofrece una mayor señal visual. Otros en cambio manifestaron percibir a todos por igual. Finalmente, se les pregunto si harían uso de estos dispositivos en el centro de educación, a lo cual, 3 educadores los utilizarían para realizar tareas puntuales como poner música, preguntar la hora o el tiempo. Sin embargo, manifestaron junto al resto de educadores que mientras existan barreras en la frase de activación descartarían su uso.

5.4 Sesiones de participantes con perfil tecnológico

La realización de las sesiones con los participantes de la Escuela Politécnica Superior, se dieron a cabo en los meses de marzo, abril y mayo. Las mismas se daban sin un orden específico debido a su disponibilidad de tiempo para lo cual se utilizó la herramienta Doodle con lo cual se coordinaron las sesiones.

5.4.1 Inicio de cada sesión

En el inicio de las sesiones se entregó un documento a los participantes en el cual se explicó que realizarían un grupo de tareas con cada asistente. El documento estaba formado por 3 hojas, una para cada asistente con su respectiva frase de activación. Se les indicó que para realizar cada petición debían usar la frase de activación que estaba escrita en la parte superior de la hoja. Se ejecutó una tarea de prueba con cada asistente para que supieran como era la forma de interacción con los dispositivos.

5.4.2 Durante la sesión

Durante la ejecución del primer grupo de tareas en donde se debía controlar las luces y el ventilador no existieron problemas de activación ni de ejecución por parte de los participantes, sin embargo, en las sesiones realizadas por los participantes 7 y 19 en donde la tarea era llevada a cabo por el asistente de voz Siri, el cual devolvía un mensaje indicando que había un problema con el dispositivo por lo cual no era posible ejecutar la tarea. De igual manera no existieron inconvenientes al ejecutar la tarea sobre poner el temporizador y detenerlo. De igual manera, en el segundo grupo de tareas, no existieron problemas al realizar las peticiones sobre la hora. Por otro lado, los participantes notaron cierta demora en la respuesta al realizar la petición del tiempo, con el altavoz HomePod. Por otro lado, cuando se realizaba la petición de cuanta falta para Semana Santa, cada asistente daba información de manera diferente. Finalmente, en el tercer grupo de tareas en donde los participantes podían realizar preguntas sobre temas de interés personal no hubo mayores problemas al responder las peticiones salvo en algunos casos en donde se preguntaba a los asistentes sobre lugares donde caminar o andar en bicicleta.

5.4.3 Entrevista al finalizar la sesión

Luego de la sesión con los participantes de la Escuela Politécnica Superior de la UAM, se realizó una entrevista sobre su experiencia con los altavoces. Con lo cual, con las preguntas mostradas en el Anexo B, se obtuvo la siguiente información:

En general manifiestan que son fáciles de usar y que han disfrutado de las sesiones. Al preguntarles sobre su opinión de cada asistente han respondido que, cumplen su función como asistentes puesto que interpretan y responden bien. Sin embargo, manifestaron que el asistente de Google es el que presenta mayor dificultad para interactuar debido a que en su opinión es menos intuitivo, difícil de pronunciar y además es un nombre extraño a utilizar en una conversación. Por otro lado, algunos consideran que las respuestas obtenidas por los asistentes en general son muy artificiales. Indican que en un principio no es complejo el apoyo de personal técnico puesto que son intuitivos. Al preguntarles sobre que asistente es más intuitivo, consideran que los dispositivos que contienen a Alexa y Google Assistant son más intuitivos por las luces presentadas. Finalmente, la opinión sobre el asistente que usarían en el futuro es dividida debido a que, consideran que estos dispositivos no ofrecen privacidad puesto que están escuchando todo el tiempo y sus funcionalidades son limitadas. Otros manifiestan que los usarían para realizar búsquedas rápidas, reproducir música o preguntar por direcciones.

6 Resultados

Al realizar el análisis de las grabaciones realizadas sobre las sesiones, se descartó al participante 5 de todo el estudio debido a que, la barrera de comunicación que presentan es demasiado severa por lo que al realizar las sesiones no pudo activar ninguno de los altavoces ni ejecutar ninguna de las tareas.

Durante la sesión realizadas con el altavoz HomePod, el participante 8, realizó únicamente 3 tareas del primer grupo de actividades debido a la frustración generada. Esto se debió a que su pronunciación es pausada con lo cual no pudo activar el altavoz. De igual manera el participante 4 realizó 2 tareas, debido a que en el momento de la sesión indicó que no deseaba encender el ventilador. Finalmente, el participante 1 y 4 no realizaron las tareas de temporizador puesto que los educadores indicaron que no se usaran esas tareas con ellos. Por otro lado, en el segundo grupo de tareas, los educadores indicaron que el participante 22 no sabe sumar por lo que esta tarea no se realizó con este usuario.

En las sesiones realizadas con el altavoz Google Home, en el primer grupo de tareas los participantes 22 y 23 no realizaron la petición referente al temporizador por indicación de los educadores. Así mismo, el participante 23 no realizó las tareas referentes al ventilador puesto que, el educador intervino al notar frustración cuando intentó activar el altavoz y no lo logró. En el segundo grupo de tareas, los participantes 12 y 17 no realizaron las tareas referentes a la petición de “¿Cuánto falta para semana santa?” por indicación del educador. De igual manera los educadores de los participantes 12, 13, 14 y 17 indicaron que sus estudiantes no realizarían la tarea de la suma debido a que informaron que no saben sumar o realizan este tipo de operaciones con una estructura especial para su entendimiento.

En las tareas realizadas con Amazon Echo, la frase de activación usada para la tarea de programar el temporizador no fue realizada por los participantes 9, 12, 13 y 16 por indicación del educador. En otro caso el participante 17 realizó un solo intento de activar el altavoz en la tarea de encender la luz puesto que el educador notó frustración del participante por lo que intervino en la sesión. Finalmente se brindó apoyo por parte del observador al participante 10 en 3 tareas de la primera parte debido a que, el participante pronunciaba la frase de activación por sílabas. Para el segundo grupo de tareas el participante 4 no realizó la tarea de la suma por indicación del educador. Por otro lado, el participante 8 no realizó ninguna tarea debido a la frustración causada por el primer asistente que usó puesto que no pudo ejecutar las frases de activación.

A continuación, se realizó el análisis explicado en la metodología con lo cual se han obtenido los siguientes resultados:

6.1 Frase de activación

Como primer punto de análisis, se toma en cuenta la frase de activación usada para que los asistentes de voz comiencen a recibir peticiones. El análisis se ha realizado por asistente de voz, para lo cual, en la Tabla 8, se muestra la cantidad total de tareas realizadas por cada asistente. Como se ha mencionado en los puntos anteriores, el número de tareas es variable debido a las indicaciones realizadas por los educadores. Sin embargo, por distintas situaciones que surgieron durante las sesiones, algunas tareas no fueron realizadas durante las sesiones debido a frustración o negación por parte de los participantes.

Altavoz	G. Tareas 1	G. Tareas 2	G. Tareas 3	Total Tareas
Homepod	31	32	41	104
Google Home	47	29	38	114
Amazon Echo	37	23	46	106

Tabla 8: Numero de tareas realizadas por altavoz.

6.1.1 Activación usando “Oye Siri”

El número de tareas realizadas por los participantes del centro de educación ALENTA con el altavoz HomePod han sido 104. Este número está compuesto por todas las tareas realizadas de cada grupo de actividades y cada usuario. Con lo cual, se observa en la Figura 6.1, el porcentaje de activaciones realizadas en el primer intento por los participantes es del 71% mientras que un 23% han sido realizadas por el observador, solo un 6% es realizado en el segundo y tercer intento. Esta activación por parte del observador se ha dado en casos en concreto en donde los usuarios han presentado problemas al enlazar las 2 palabras de la frase de activación puesto que realizan una pausa al pronunciarlo. En el caso de los participantes 9 y 12 realizaban la petición llamando al asistente con una sola palabra, en este caso Siri u Oye. Por otro lado, el participante 1 no tenía problemas de enlazar las palabras de la frase de activación, sin embargo, tuvo dificultad al pronunciar Siri puesto que presenta una barrera de comunicación con lo cual vocalizaba algo parecido a “Sisi” o “Isi”.

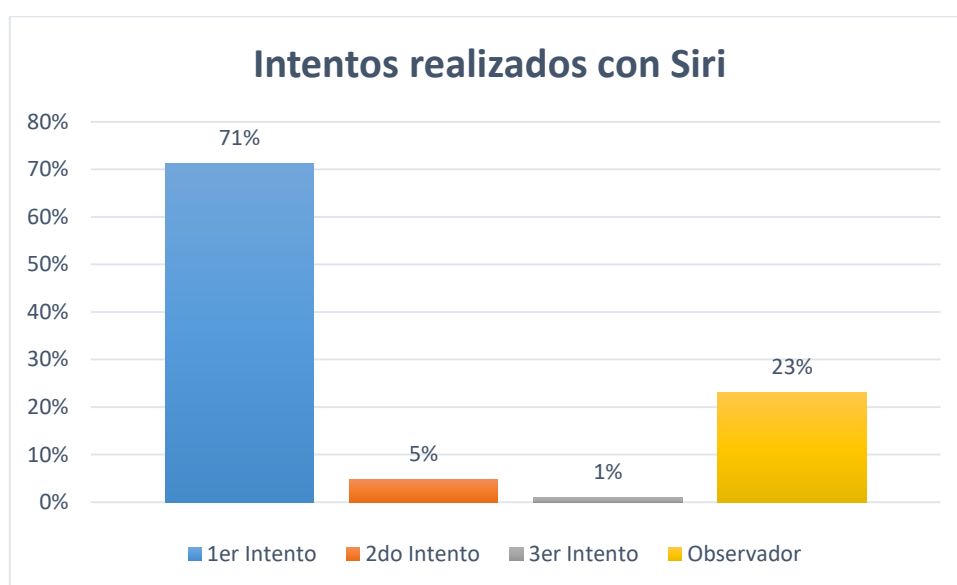


Figura 6.1: Intentos realizados con Siri

6.1.2 Activación usando “Ok Google”

Con el asistente de voz de Google se realizaron 114 tareas por los diferentes participantes del centro de educación ALENTA. Al realizar el análisis sobre la frase de activación de Google Assistant, en la Figura 6.2, se observa que, el 51% de frases de activación realizadas por los participantes han sido ejecutadas en el primer intento mientras que en el segundo y tercer intento se realizaron con un 14% y 5% respectivamente. Sin embargo, un 30% de activaciones fueron realizadas por el observador. Esto se ha debido a que, durante las sesiones los participantes han experimentado dificultad para pronunciar la frase de activación de este asistente puesto que la misma está formada por 2 palabras las cuales son difíciles de pronunciar. En el caso del participante 19 no tuvo problemas en ejecutar las frases de activación con los asistentes Alexa y Siri, sin embargo, cuando intentaba realizar esta acción con Google Assistant se le dificultó por problemas de vocalización con la palabra “Google” puesto que tartamudeaba al tratar de pronunciarla.

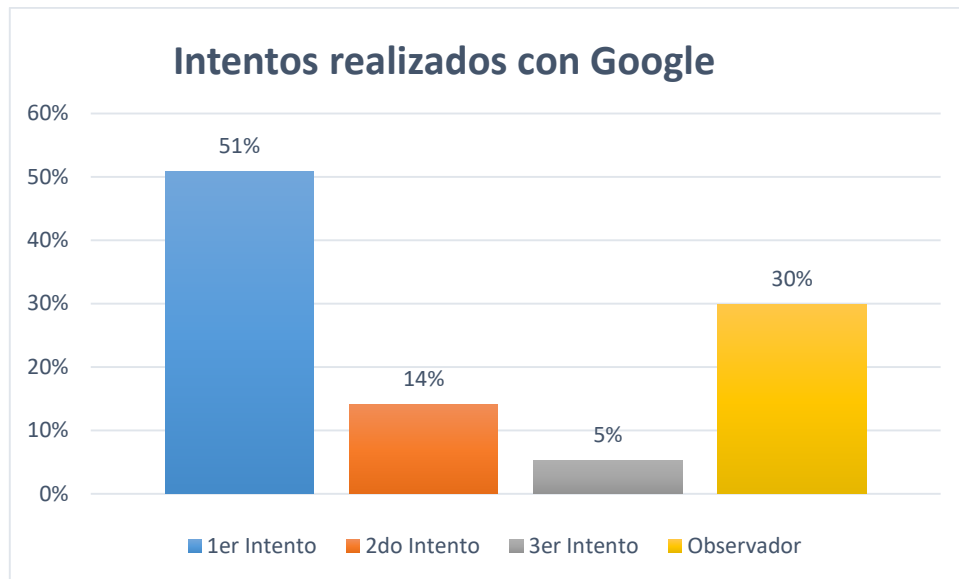


Figura 6.2: Intentos realizados con Google

Durante las sesiones, se ha escuchado que los participantes pronuncian de diferentes maneras la frase de activación de Google Assistant. Algunos ejemplos escuchados son “Ok Buguel”, “Ok Gugue”, “Ok Bugle”, entre otros. En algunos casos estas variaciones han activado el altavoz, esto se ve reflejado en el segundo y tercer intento realizado.

6.1.3 Activación usando “Alexa”

En la Figura 6.3, se muestra que, el 45% de frases que activaron el altavoz tuvieron lugar en el primer intento. Por otro lado, un 35% de estas frases fueron activadas por el observador con lo cual, queda en evidencia que esta frase de activación dificulta la interacción con el dispositivo. A pesar de que la frase de activación del asistente de Amazon lleva una sola palabra, la fonética de “Alexa” envuelve dificultades para personas dentro de la diversidad funcional intelectual puesto que la pronunciación de esta palabra presenta cierta dificultad en relación a la fonética.

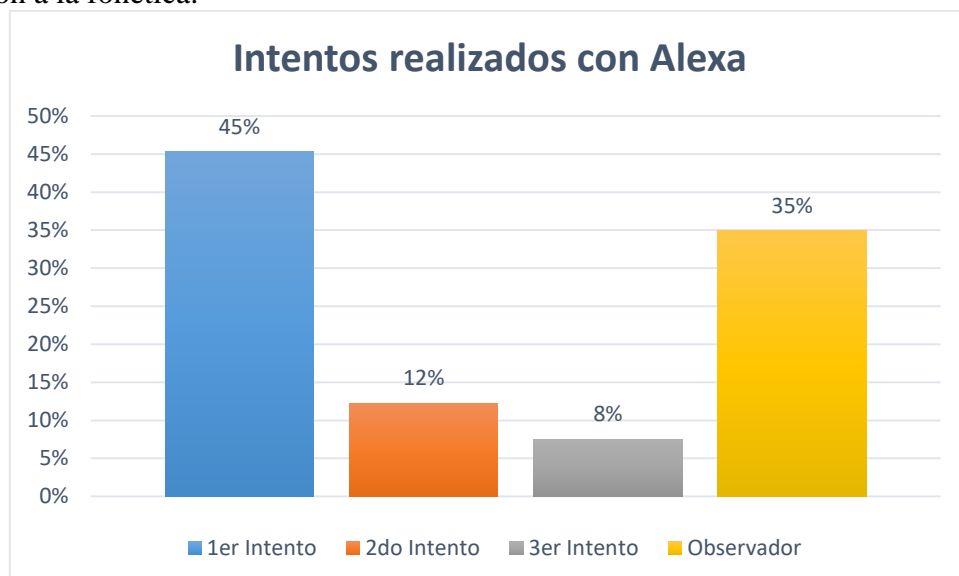


Figura 6.3: Intentos realizados con Alexa

En ciertas ocasiones, los participantes activaron los altavoces pronunciando variaciones de la frase de activación como Aleza, Alexia o Alezia. Sin embargo, estas activaciones tenían un tiempo menor de espera para la recepción de la petición. En consecuencia, aquellos

participantes realizaban un esfuerzo mayor para que el asistente los escuchase y en ocasiones al percibir su activación trataron de ejecutar las tareas, sin embargo, estas no fueron escuchadas.

6.1.4 Comparación entre comandos de activación

Como se observa en la Figura 6.4, la cantidad de activaciones por número de intento varía entre cada asistente siendo Siri el que mayor número de activaciones tiene en el primer intento, de igual manera, el que menor número de activaciones han sido realizadas por el observador. Por otro lado, al analizar los intentos realizados para activar a los asistentes Google Assistant y Alexa, se puede observar que el porcentaje es mayor a los que han sido necesarios con Siri en el segundo y tercer intento. Así mismo, el observador ha tenido que activar en mayor cantidad de veces a los asistentes Alexa y Google Assistant para que los participantes puedan continuar con la realización de las tareas.

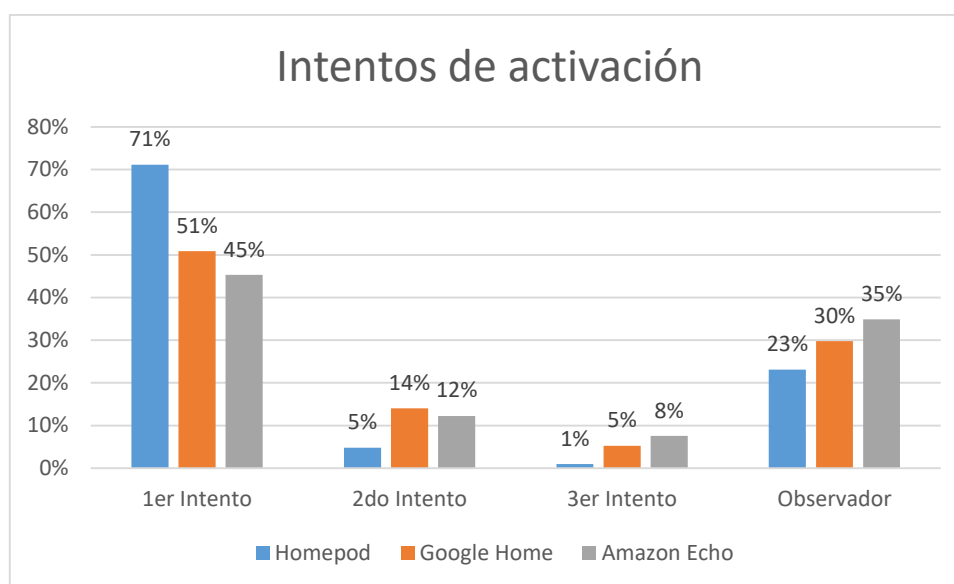


Figura 6.4: Comparativa entre frases de activación

Los educadores mencionan que, aunque pueden pronunciar las palabras, el problema reside en unirlos en frases por lo que, lo realizan de forma pausada con lo cual no logran activar los asistentes. Además, comentan que no es natural usar 2 palabras para referirse a los nombres de los asistentes.

De acuerdo a lo analizado en cada sesión de los participantes, las frases de activación de los asistentes de voz desempeñan un papel importante puesto que la dificultad de pronunciarlas genera una barrera para poder utilizarlos. En el caso de la población con diversidad funcional intelectual, esta barrera causa aún mayores dificultades puesto que la tolerancia a la frustración es baja por lo que al no poder activar estos dispositivos pierden el interés en los mismos. En concreto esta pérdida de interés se ha dado con los participantes 4, 8, 12 y 17 puesto que a medida que se realizaban las sesiones, se notó su frustración por lo cual realizaron menos tareas. Por otro lado, la tolerancia a la frustración del participante 14 es mayor puesto que realizó la sesión completa con los 3 altavoces, sin embargo, en ninguno de los casos pudo activar los asistentes, pese a ello pudo realizar ciertas tareas de las sesiones como encender o apagar la luz o al pedirle que le cuente un chiste.

6.2 Tareas realizadas

En el estudio realizado, se propone analizar:

1. Petición de tarea: El participante realizara la petición indicada al asistente. Se tomará en cuenta la formulación de la misma.
2. Respuesta obtenida: Respuesta de acorde a la petición.

El análisis de tarea realizada toma en cuenta a los participantes que, pese a que no han podido activar los asistentes han ejecutado las tareas. Esto se hace debido a que, en el análisis anterior se encontró limitaciones al activar los altavoces debido a la complejidad que presentan ya sea por fonemas complicados que presente la frase o porque está formada por 2 palabras que necesitan ser pronunciadas de manera continua y sin pausa. A continuación, se realiza el análisis de los resultados obtenidos por altavoz. Presentando las tareas en donde se han evidenciado un mayor número de problemas.

6.2.1 Tareas realizadas con Siri

La Figura 6.5, muestra el primer grupo de tareas realizadas con el altavoz HomePod con el cual se observa que al realizar la tarea de encender la luz, un 67% han realizado la petición correcta a ejecutar sin embargo, el asistente no ha realizado la acción en todos los casos. Esto se ha dado debido a que, pese a que el participante 2 formulo la petición de manera correcta, el asistente respondió de manera errónea a la solicitud. Así mismo, cuando el participante 6 realizo la petición de encender la luz, no se obtenía una respuesta errónea, simplemente no realizaba ninguna acción.

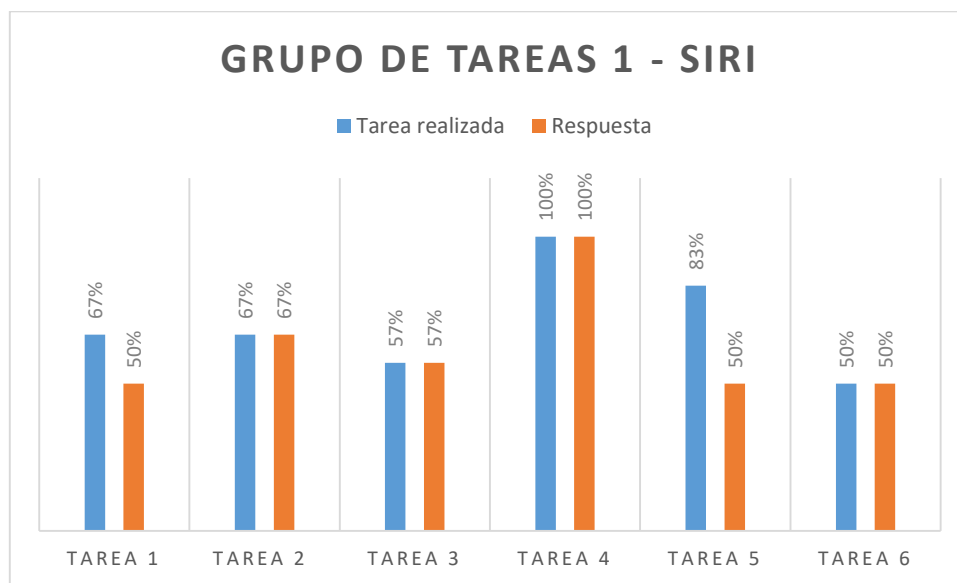


Figura 6.5: Primer grupo de tareas realizadas con Siri

En el caso de la tarea del temporizador, debido a la complejidad y longitud de la tarea, el participante 3 realizo la petición sin embargo el asistente solo entendió la primera parte por lo que solicito que se estableciera el tiempo, a lo cual el participante intento establecer el tiempo, pero el asistente no entendió la orden. Así mismo debido a la pronunciación de la palabra ventilador, los participantes que trataron de realizar esta tarea no lograron completarla puesto que pronunciaban frases parecidas a “talador” refiriéndose al ventilador con lo cual solo un 50% pudo ejecutar la misma.

La Figura 6.6 muestra que, el segundo grupo de tareas los participantes han logrado realizarlas sin mayores complicaciones. Salvo en el caso de la tarea 4, en donde la longitud de pregunta es mayor a las anteriores. El participante 19, el cual presenta una pronunciación clara, sin embargo, realiza pausas al hablar lo que produjo que el asistente tomara fragmentos de la solicitud y daba respuestas erróneas.

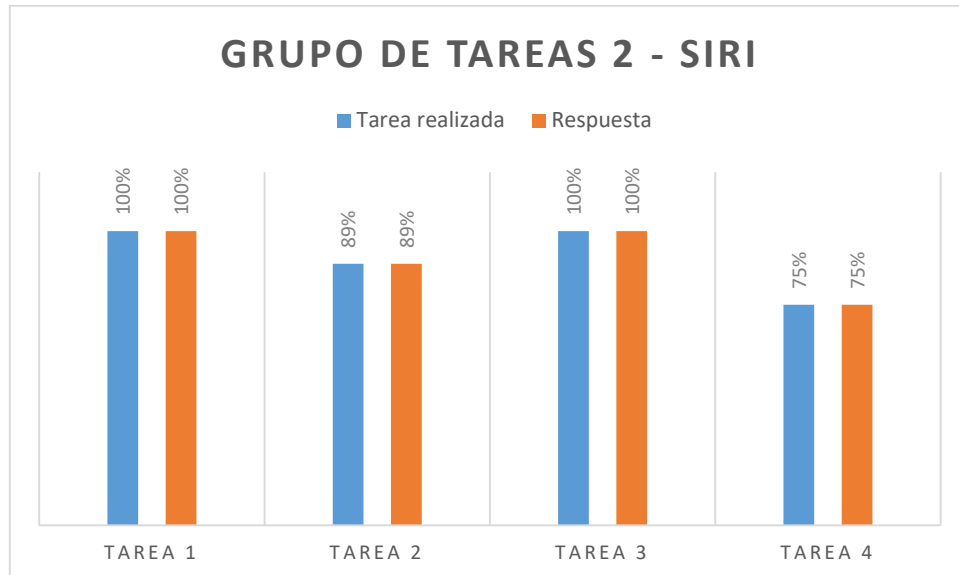


Figura 6.6 Segundo grupo de tareas realizadas con Siri

En el tercer grupo de tareas en donde se les permitió realizar a los participantes las peticiones que desearan, existieron un mayor número de errores por parte de participantes los cuales su nivel de pronunciación y comunicación era más severo puesto que realizaban peticiones incompletas e ininteligibles. Con lo cual, al no existir una estructura a seguir para realizar estas tareas, las respuestas obtenidas no será las deseadas en todos los casos.

6.2.2 Tareas realizadas con Google

Al analizar los resultados del grupo de tareas 1 realizadas con Google Assistant mostrados en la Figura 6.7, se encuentra que, al realizar la tarea sobre poner el temporizador, solo un 43% pudo realizarla. El participante 18, pese a presentar un nivel de pronunciación claro, habla rápidamente, con lo cual tuvo problemas para realizar esta tarea, debido a que el asistente tomaba solo la mitad de la petición y solicitaba que se confirme el tiempo de duración del temporizador, sin embargo, no logro culminar la tarea.

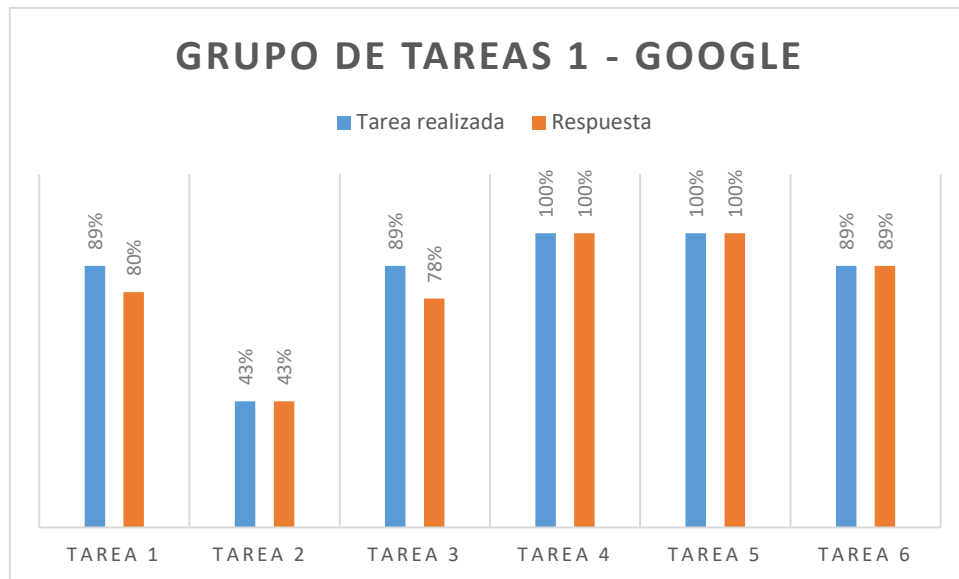


Figura 6.7: Primer grupo de tareas realizadas con Google

De igual manera, en la tarea de encender el ventilador, el participante 18, pronuncia la petición de manera interpretable, la respuesta obtenida del asistente es la definición de ventilador.

La Figura 6.8, muestra los resultados del segundo grupo de tareas en donde se aprecia que han surgido ciertos inconvenientes para ser llevadas a cabo. Cuando el participante 10 realizo la petición de obtener la hora, el asistente no retorno ningún mensaje. En otro caso el participante 12 el cual presenta dificultades de articulación al hablar, realizo la misma tarea y de igual manera que en el caso anterior no obtuvo una respuesta por parte del asistente con lo cual manifestó que ya no quería realizar esa tarea. En las tareas 3 y 4, la dificultad encontrada es por un lado la articulación que realizan y por otro lado la organización de la tarea a realizar.

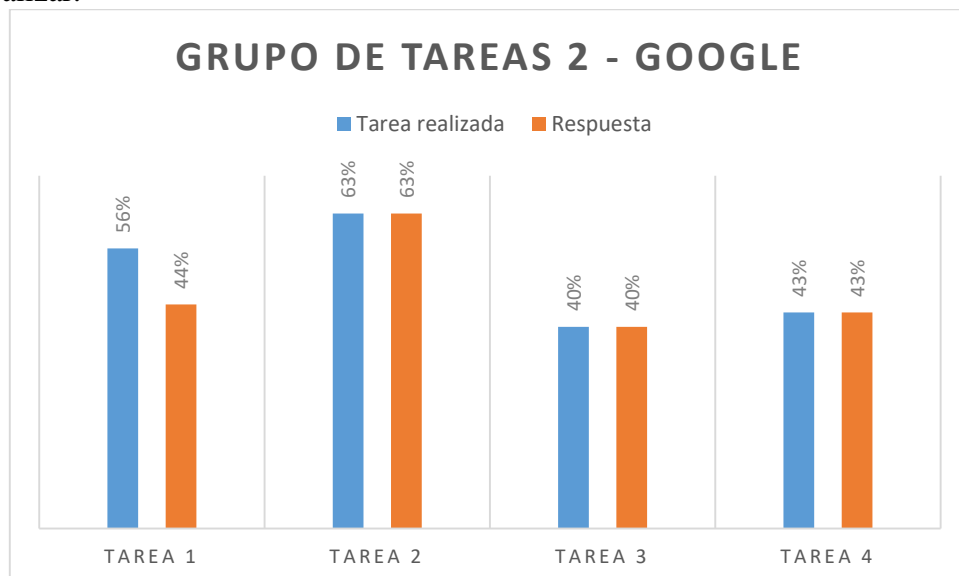


Figura 6.8: Segundo grupo de tareas realizadas con Google

En las tareas realizadas en la parte de conversación libre con Google Assistant, quienes poseen un nivel de pronunciación y comunicación alto, pueden realizar peticiones sin problema. Las peticiones más realizadas fueron acerca de pedir música, preguntar resultados de deportes o fechas de partidos de futbol. Sin embargo, para quienes presentan dificultades

articulatorias y emplean estructuras medias para comunicarse, presentan una menor cantidad de respuestas satisfactorias.

6.2.3 Tareas realizadas con Alexa

La Figura 6.9, muestra el primer grupo de tareas realizadas con el asistente Alexa. Solo un 33% de participantes pudieron realizar la tarea 1. Los participantes 10, 11 tienen una pronunciación clara, sin embargo, les cuesta articular la palabra temporizador y unirla a la petición completa con lo cual no pudieron formular la petición. En la tarea 4 en la cual se solicita que se detenga el temporizador, el participante 11 realiza la petición de detención del asistente sin embargo el asistente continuó ejecutando el temporizador. Por otro lado, en las tareas de encender y apagar el ventilador las cuales corresponden a la tarea 3 y 6, los participantes formularon de manera correcta la petición, sin embargo, la articulación de los mismos dificultó que el asistente ejecutara la petición.

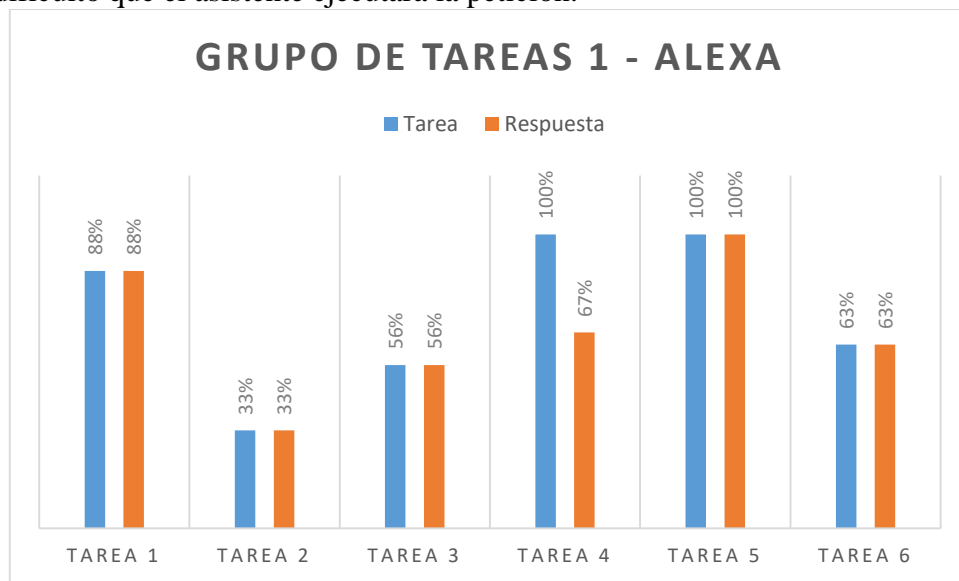


Figura 6.9: Primer grupo de tareas realizadas con Alexa

La Figura 6.10, muestra un resumen de las tareas ejecutadas por los usuarios, en donde se observa que, la comprensión del asistente por parte de la formulación y pronunciación de los participantes es la correcta. En casos en donde se observa un porcentaje menor son en las preguntas referentes a realizar una suma y preguntar sobre una fecha en concreto.

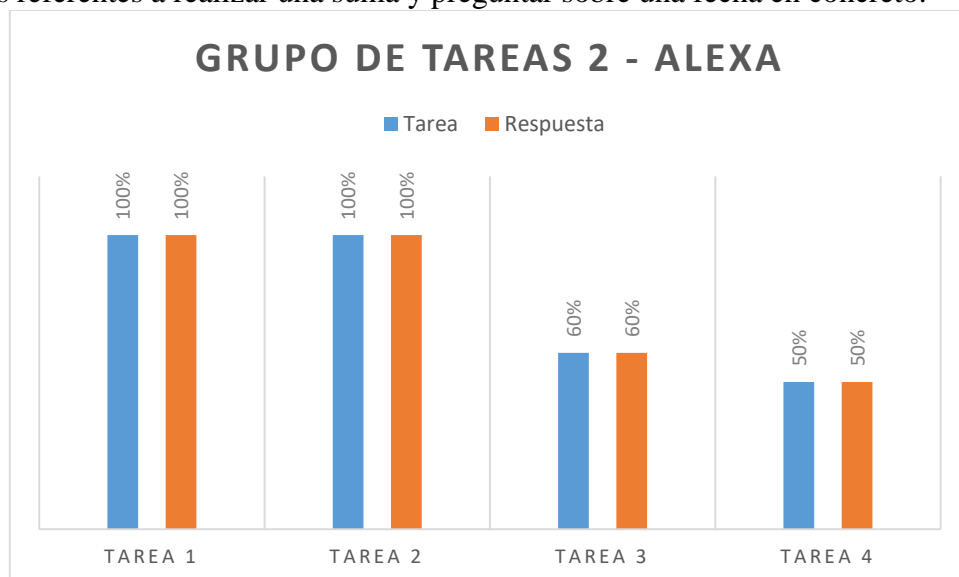


Figura 6.10: Segundo grupo de tareas realizadas con Alexa

En el tercer grupo de tareas en el cual, los participantes realizaron las peticiones de acuerdo a intereses propios. Se analizan las respuestas obtenidas por los asistentes. En el caso del participante 21 al realizar la pregunta “Cuando es mi cumpleaños”, el asistente responde que no lo sabe pero que le desea un día fantástico. El participante 20 realizó varias preguntas referentes a fútbol pese a que articula bien el asistente expresa que no las tiene. En otros casos en donde el grado de pronuncia es ininteligible, la formulación de las tareas ha sido interpretable sin embargo no han sido capaces de ser respondidas adecuadamente.

7 Discusión

7.1 Frase de activación

En el punto anterior se ha revisado los distintos errores cometidos al pronunciar las frases de activación. Quienes participaron en el estudio, presentan diferentes niveles de pronunciación en donde sus dificultades articulatorias son particulares en cada uno.

De acuerdo a lo observado durante las sesiones de prueba, la frase de activación “Alexa” presenta un fonema difícil de pronunciar debido a la presencia de la letra x en el nombre. Por otro lado, al analizar la frase de activación “Ok Google” y “Oye Siri” se evidencia otro problema que enfrenta esta población, la cual es la memorización con lo cual exige que pronuncien la palabra y recuerden cual es la que sigue. Adicional a ello, aquellos participantes que realizan pausas entre palabras no consiguen activar los asistentes. En el caso de Google, es una palabra en inglés y no es común usarla en una conversación con lo cual, dificulta la interacción con este altavoz.

Al analizar los problemas encontrados se considera que, la frase de activación ideal para un asistente debería hacer uso de una sola palabra, la misma que debe ser fácil de pronunciar y no haga uso de fonemas de difícil articulación. Así mismo, que la percepción de la frase de activación sea capaz de ser activada en los casos de personas con una pronunciación pausada.

7.2 Tareas realizadas

El éxito de que las tareas realizadas por los participantes, se basa en la longitud de petición y uso de palabras conocidas por los mismos. En los casos en donde se ha pedido que realicen las tareas que usaban el temporizador o ventilador es en donde existieron mayores inconvenientes puesto que por el lado del temporizador, la dificultad se basa en la palabra y seguidamente indicar el tiempo de duración. En el caso en donde se usaron dispositivos para controlar las luces y el ventilador, existieron problemas de comunicación, pese a que previo al inicio de cada sesión se verificaba el correcto funcionamiento de la luz y el ventilador. En algunas ocasiones durante las sesiones estos no funcionaban con lo cual, aunque indican que están diseñados para funcionar mediante el uso de asistentes de voz, esta tecnología aun no es óptima.

En el grupo de tareas en donde se permitió libertad de preguntar lo que los participantes deseaban se evidencio también problemas debido a que al sugerirles lo que podrían preguntar, lo hacen de acuerdo a su pronunciación y organización del lenguaje. Un ejemplo de ello fue con el participante 11 del grupo 2, el cual realizo la petición “*Quiero que me pongas Amaia Montero, 100 metros*” y el asistente responde “*Lo siento, pero no he encontrado lo que estás buscando*”, tras recibir esta respuesta el usuario se nota descontento por lo que el evaluador realiza la petición de la siguiente manera “*Ponme la canción de Amaia Montero, 100 metros*” y el asistente comienza a reproducir la canción solicitada. El usuario piensa un momento y comenta que es necesario realizar la petición lentamente, pese a que la velocidad de habla del mismo es normal.

Por otro lado, en preguntas interactivas como cuéntame un chiste o cántame una canción, las respuestas obtenidas varían entre los asistentes. De acuerdo a las sesiones en donde se realizaron estas peticiones se nota en donde, este tipo de respuestas son más cálidas con el usuario. A esto se suma una crítica realizada por uno de los educadores cuando se usó estas preguntas con el asistente de voz Siri en la cual señalo que las respuestas dadas son las menos amigables y que no inspiran a sus alumnos.

7.3 Percepción de activación

De acuerdo a las opiniones recolectadas durante la entrevista tanto por parte de educadores del centro de educación especial como por parte de participantes con perfil tecnológico, manifiestan que, las señales visuales que muestran los altavoces al ser activados es diferente entre cada dispositivo puesto que, en su mayoría indica que, el anillo de luz presentado por el altavoz Amazon Echo es más fácil de detectar puesto a que, aunque los altavoces de Google y Apple también muestran estas señales, el contraste con el color del dispositivo provoca una limitación en la percepción de activación. Esto se ha podido constatar al revisar las grabaciones de las sesiones puesto que la ubicación de esta luz se encuentra en un ángulo de fácil visualización, Mientras que en los otros dispositivos es necesario prestar atención en la parte superior del altavoz.

7.4 Guías de diseño propuesta

- Frase de activación: Debe tomar en cuenta los fonemas que pueden presentar problemas al articular las palabras. En lo posible debe hacer uso de una sola palabra. Además, la palabra clave debe ser representativa.
- Orden de frase de activación: Se considera que la frase de activación tome en cuenta las peticiones realizadas antes de la pronunciación de la frase puesto que la organización de la estructura del lenguaje en el entorno de la diversidad funcional intelectual varía de persona a persona.
- Tiempo de espera: Se considera que, el tiempo que el asistente espera para recibir una petición debe ser variable debido a que, quienes forman parte de la población considerada con diversidad funcional intelectual presentan problemas en la comunicación y como se constató en las sesiones y por comentarios realizados por los educadores, les toma mayor tiempo formular y realizar una petición.
- Señales visuales de percepción de activación: Refiriéndose a la facilidad de distinguir que se ha activado el asistente y a la ubicación de estas señales. Estas deben situarse de tal forma que sean de fácil visualización sin importar la distancia y ubicación en la que se encuentren los usuarios. Por otro lado, los colores deben ser fácilmente distinguibles puesto que si el color de estos dispositivos, lo esencial es que contrasten con los mismos.
- Respuestas por defecto: En el caso de que una petición no haya sido entendida, se retorne obligatoriamente una retroalimentación. Considera variar respuestas por defecto.
- Retroalimentación de peticiones inconclusas: En el caso de que, la orden dada no genere ninguna respuesta, el asistente debe ofrecer retroalimentación haciendo uso de las palabras usadas con la petición.
- Longitud de respuesta: La longitud de respuesta no debe extenderse demasiado. En el caso de la pregunta realizada sobre el tiempo, la información obtenida por los asistentes es demasiado completa. Para quienes estén incluidos en la diversidad funcional intelectual, esta cantidad de información puede ser difícil de asimilar puesto que se especifica cuáles serán las temperaturas máximas y mínimas

8 Conclusiones y Trabajo Futuro

Al finalizar el estudio se puede concluir que, existen algunas barreras que limitan el uso de asistentes de voz por la población considerada en el entorno de la diversidad funcional intelectual. Una de estas barreras se presenta desde la invocación a los asistentes puesto que no se ha tomado en consideración que, esta población presenta un desarrollo de lenguaje reducido y diverso de persona en persona. Con lo cual, cuando han intentado comunicarse con los asistentes, no han podido hacerlo debido a que, las palabras usadas presentan consideraciones importantes en su fonética. Del mismo modo, los asistentes que hacen uso de 2 palabras en la frase de activación, requieren que se realice sin pausas, puesto que, al no cumplir este requerimiento, los asistentes no se activan. Esta situación se repite, si la frase de activación no se pronuncia de forma clara.

Las ventajas que ofrecen estos dispositivos está limitada a tareas en donde la petición es corta puesto que, en las sesiones realizadas, se evidencio que las tareas propuestas en las cuales se hacía uso de frases largas y que requieren una estructuración lógica para su apropiada ejecución. Adicional a lo antes mencionado, el tiempo de escucha otorgado por los asistentes es limitado puesto que, si se realizan pausas en medio de la petición, el fragmento pronunciado será usado por el asistente, con lo cual, de acuerdo a la complejidad de la pregunta, pueden o no responder de manera correcta.

La metodología utilizada en el estudio permitió probar las funciones propuestas con todos los asistentes de manera equitativa. De la misma manera, evito una sobrecarga y posible frustración y abandono del estudio por parte de los participantes puesto que, desde un principio uno de los riesgos que se corría al realizar el estudio era la negación a continuar dentro del mismo si los participantes se sentían frustrados.

Finalmente, se ha llegado a la conclusión de que, por el momento, no existe un asistente de voz idóneo para ser usado en el entorno de la diversidad funcional intelectual por las barreras que limitan su uso. Con lo cual, el presente estudio propone unas guías de diseño las cuales toman en consideración los problemas encontrados al usarse con los participantes. De igual manera la guía propuesta se apoya en la información recolectada por los educadores que estuvieron presentes durante las sesiones y que pudieron brindar una visión global de los requerimientos que tiene esta población para ser incluida en el uso de tecnología.

Trabajo Futuro

Como trabajo futuro se planea realizar consideraciones adicionales analizando las sesiones grabadas. Algunos puntos descritos a continuación han sido nombrados en la discusión.

- Medir tiempos de espera para petición: El tiempo de espera para petición es el intervalo que brindan los asistentes desde que son invocados hasta la recepción de una tarea a realizar.
- Analizar la percepción de activación: Como se ha recalcado antes, quienes están considerados dentro de la diversidad funcional intelectual hacen uso de apoyos visuales a lo largo de su formación por lo cual, las señales que muestran los altavoces inteligentes deben captar la atención de manera sencilla del usuario con lo cual en trabajos futuros se podría analizar esta percepción.
- Capacidad de aprendizaje: Se planea medir si existe un aprendizaje en el uso de los asistentes ya sea por el lado de frase de activación como por petición realizada.

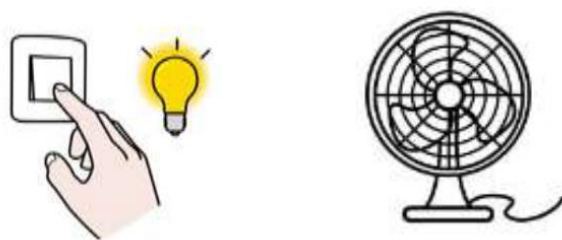
Referencias

- [1] A. Shree and P. C. Shukla, "Intellectual disability: Definition, classification, causes and characteristics," *Learn. Community-An Int. J. Educ. Soc. Dev.*, vol. 7, no. 1, pp. 9–20, 2016.
- [2] A. P. Association and others, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.
- [3] M. Hoy, "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants," *Med. Ref. Serv. Q.*, vol. 37, pp. 81–88, 2018.
- [4] R. C. Simpson, *Computer access for people with disabilities a human factors approach*. Boca Raton; London; New York: CRC Press, 2013.
- [5] M. Baldauf, R. Bösch, C. Frei, F. Hautle, and M. Jenny, "Exploring Requirements and Opportunities of Conversational User Interfaces for the Cognitively Impaired," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 2018, pp. 119–126.
- [6] C. Carroll, C. Chiodo, A. X. Lin, M. Nidever, and J. Prathipati, "Robin: Enabling Independence For Individuals With Cognitive Disabilities Using Voice Assistive Technology," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017, pp. 46–53.
- [7] T. Wallace and J. Morris, "Identifying Barriers to Usability: Smart Speaker Testing by Military Veterans with Mild Brain Injury and PTSD," in *Breaking Down Barriers*, 2018, pp. 113–122.
- [8] S. S. Balasuriya, L. Sitbon, A. A. Bayor, M. Hoogstrate, and M. Brereton, "Use of Voice Activated Interfaces by People with Intellectual Disability," in *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, 2018, pp. 102–112.
- [9] C. Myers, A. Furqan, J. Nebolsky, K. Caro, and J. Zhu, "Patterns for How Users Overcome Obstacles in Voice User Interfaces," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, p. 6:1--6:7.
- [10] F. Ballati, F. Corno, and L. De Russis, "Assessing Virtual Assistant Capabilities with Italian Dysarthric Speech," in *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 2018, pp. 93–101.
- [11] S. Melvin Felix, S. Kumar, and A. Veeramuthu, "A Smart Personal AI Assistant for Visually Impaired People," 2018, pp. 1245–1250.
- [12] P. Bose, A. Malpethak, U. Bansal, and A. Harsola, "Digital assistant for the blind," in *2017 2nd International Conference for Convergence in Technology, I2CT 2017*, 2017, vol. 2017–Janua, pp. 1250–1253.
- [13] R. Chen, Z. Tian, H. Liu, F. Zhao, S. Zhang, and H. Liu, "Construction of a voice driven life assistant system for visually impaired people," in *2018 International Conference on Artificial Intelligence and Big Data, ICAIBD 2018*, 2018, pp. 87–92.
- [14] J. Lewis, "Google Home," *Computer Act!ve*, 2017. [Online]. Available: https://store.google.com/es/product/google_home_specs.
- [15] J. Nielsen, *Usability Engineering*. Amsterdam[etc.: AP Professional-Morgan Kaufmann, 1993.
- [16] V. L. Anderson, *Design of experiments : a realistic approach*. New York: M. Dekker, 1974.

Anexos

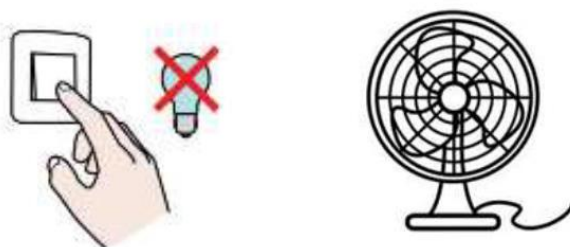
Anexo A. Pictogramas usados en las sesiones.

Las siguientes figuras muestran los pictogramas utilizados durante las sesiones con los participantes del centro de educación especial ALENTA



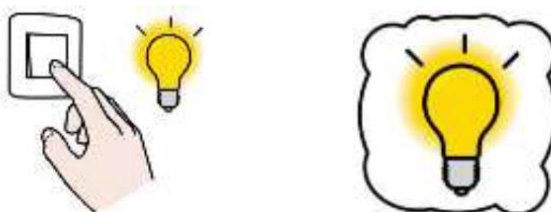
ENCIENDE EL VENTILADOR

Figura 8.1: Pictograma de encender el ventilador



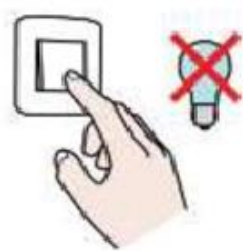
APAGA EL VENTILADOR

Figura 8.2: Pictograma Apaga el ventilador



ENCIENDE LA LUZ

Figura 8.3: Pictograma Enciende la luz



APAGA

LA LUZ

Figura 8.4: Pictograma Apaga la luz



¿QUÉ

TIEMPO

HACE?

Figura 8.5: Pictograma de pregunta ¿Qué tiempo hace?



CUÉNTAME

UN

CUENTO

Figura 8.6: Pictograma Cuéntame un cuento



CÁNTAME UNA CANCIÓN

Figura 8.7: Pictograma Cántame una canción

Anexo B. Preguntas de entrevista

Las preguntas utilizadas para la entrevista después de la sesión con los participantes de la Escuela Politécnica Superior de la UAM se utilizaron con el fin de obtener información sobre limitaciones que hayan encontrado los participantes sobre las frases de activación, ejecución y respuestas obtenidas.

1. ¿Te ha gustado usar los asistentes? Sí___ No___ ¿Por qué?
2. Qué opina del asistente de voz Siri
3. Qué opina del asistente de voz Alexa
4. Qué opina del asistente de voz Google Assistant
5. ¿Cuál de los asistentes te ha entendido mejor y por qué?
6. ¿Cuál de los asistentes ha entendido peor y por qué?
7. ¿Cuál de los asistentes da respuestas que se entienden mejor y por qué?
8. ¿Cuál de los asistentes da respuestas más difíciles de entender y por qué?
9. ¿Cree Ud. que es necesario el apoyo de personal técnico para la utilización de los asistentes? ¿Por qué?

10. ¿Con cuál de los asistentes existe mayor dificultad para interactuar y por qué?
(refiriéndose a la frase de activación o a las preguntas realizadas)
11. ¿Qué asistente es más intuitivo para saber que se le está hablando y por qué?
12. ¿Qué asistente de voz usarías en el futuro y para qué lo usarías?